ENABLING THE COLLECTIVE BRAIN FOR ORGANIZATIONS:

A QUICKSTART IN MANAGEMENT SOFTWARE SKILLS

Fons Wijnhoven

Source of image on frontpage: http://io9.com/5962914/the-emerging-science-of-collective-intelligence--and-the-rise-of-the-global-brain

Enabling the collective brain for organizations:

A quickstart in management software skills

First edition

Fons Wijnhoven

Enschede, Netherlands, September, 2013

First edition published Enschede, September 2013 as PDF for International Business Adminunistration Program

Published as digital book by University of Twente, Faculty Management and Governance, Department of Industrial Engineering and Business Information Systems.

Distributed by doc.utwente.nl.

Contents

Foreword and acknowledgements

In 2009, I completed an introduction to information management from the perspective of management studies and social sciences (Wijnhoven, 2009a). Thus the book was intended not to be technical. The focus was on information as a human and social asset used for communication, decision making and management. This required an analysis of the variety of the concept of information as a human and social phenomenon. Happily enough, an extensive amount of literature rooted in the philosophy of knowing exists, well summarized and introduced by C.W. Churchman (Churchman, 1971) and his students Mason and Mitroff (Mason & Mitroff, 1973), not for the purpose of analysis and philosophy per se but for practical purposes. Regarding the practical use, because of the extensive options that information technology offers us nowadays, the mentioned classical works are deficient, and so I wrote a book that linked fundamental thinking to modern technologies. Unfortunately, my assumption that students will easily recognize the practical implications and use technologies via publicly available free tutorials showed to be not correct and therefore I started writing this book with the help of many students and colleagues.

I am especially grateful to Matthieu van der Heijden, who commented on Chapter 3 and provided extensive useful material, and Maria Iacob, who delivered many insight regarding BPMN, Bizagi and an instructive case for chapter 4.

Many students have used previous versions of this book in the last two years, and commented on it. Especially I want to mention students Loes Brilman, Marten ten Kleij, Ellen Tolsma, Wendy Veldhuis, Martijn van der Wal, Anne Rietberg and Elise Eshuis for this. Joris Sibenius Trip and Hardwin Spenkelink delivered a very major re-write of a previous draft version of this book.

Thanks to all colleagues and students who helped by their feedback. The sole responsibility for this text however is with the author.

1. Introduction

Information consists of diverse representations of reality. These representations may be data, models, pictures and text. The realities that are represented may be about physical objects and events, but they also may be about more difficult to grasp feelings, opinions, beliefs, and methods for problem solving. In organizational contexts, we encounter these information types, and they all require different ways of management and software tools to help us on this. This book is written for understanding the link between different information concepts -as explained by Churchman (Churchman, 1971), Mason and Mitroff (Mason & Mitroff, 1973) and Wijnhoven (Wijnhoven, 2009a)- and practical information management tools so that you learn to manage and use different sorts of business information with modern software tools. These information management concepts have their roots in Lockean, Leibnizian, Kantian, Hegelian and pragmatic philosophies, as summarized in table 1.1.

Approach	**Definition of information**	**Management challenge**
Lockean empiricism	Data, facts, and figures that are supposed to correctly represent reality	Managing data about different entities for multiple users
Leibnizian rationalism	Causal insights about phenomena and how changes impact on other phenomena.	Using computer reasoning power to optimize decisions
Kantian epistemology	A perspective or view on some part of reality. Multiple views are mostly needed to realize a complete picture of a reality.	Representing organizational realities from multiple perspectives and integrating these.
Hegelian subjectivism	Subjective insights and expressions of beliefs to influence others and as input to dialogues for finding useful syntheses.	Representing subjective and political information and supporting their use in debates.
Singerian pragmatism	Ideas, information and methods for solving multi-dimensional problems	Using information for solving multidisciplinary problems in collaborative efforts
Organizational context	Information and information technology as an organizational asset	Organizing information resources for organizational needs and opportunities

Table 1.1: Information approaches

All these approaches exist in some way among people and in organizational settings, and they all have some merits. Churchman also names these approaches "inquiring systems", which he defines as purposeful systems for producing knowledge. We do not believe, as some philosophers of knowledge argue, that one type of information is superior to another or that they are incommensurable. We would rather state -following pragmatism (Churchman, 1971; Malachowski, 2010)- that their relevance depends on the situation and thus that we therefore need to know them all; we will be confronted with all of them in different contexts. The context on which this book focuses is the context of organizations. Thus we add an additional organizational contextual approach to the five inquiring systems. This organizational contextual approach links information to organizational strategic and operational decisions and work processes. It also brings all IT resources aligned with an organization's strategy and the whole collection of financial and technical means by which information management can be more effective and far-reaching than individuals could achieve themselves, which is the core idea of an organization as a synergetic cooperative system (Barnard, 1968). When information and information technology are integrated with an organization's need and well organized, they are not just interesting phenomena, but organizational assets that have to be well managed. Even more interesting, they could enabling information and knowledge sharing and collective decision making and collaboration. In this sense it becomes the enabler of an organizational collective brain. This idea of a collective brain will not be explained here further, but we use it as a challenging idea for exploration in this book. For exploring this collective brain, we believe that thinking about it has to go together with trying it out. Thus this book wants to deliver hands-on skills for information management after which the reader is challenged to think about what has been learned for the enablement of this collective brain in organizations.

As a generic approach to information management, we follow a method of informing which starts with perceived problems and questions and ends with answers, solutions and capabilities. This is a pragmatic approach to information (Churchman, 1971; Malachowski, 2010). The problem solving and design logic that we propose resembles design science (also see (Hevner, March, Park, & Ram, 2004; Peffers, Tuunanen, Rothenberger, & Chatterjee, 2007)). Between these two ends, there is (1) scoping of the area of interest, (2) analysis and modeling of the problem situation, (3) design of a solution, and (4) the realization of an informing solution using software and human skills (see Figure 1.1).

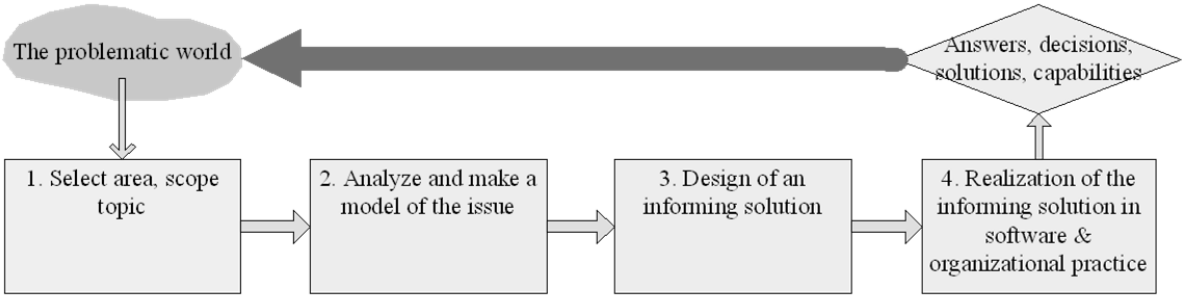


Figure 1.1: A generic model of informing and information management (Wijnhoven, 2009a)

To clarify the concepts and methods, we use the fictitious case of a Grand Café as a running case. Grand Café The Palace is a restaurant and café. They can host a maximum of 175 people. To serve its guests, Grand Café has hired staff. The staff is located in 4 departments (see table 1.2):

Department	Employees and jobs
Kitchen	o 5 cooks (full-time) o 8 assistant cooks (part-time). o 4 students (part-time). Washing the dishes.
Service	o 6 waiters (full-time) o 12 waiters (part-time) o 4 waiters on call. For events or very busy days.
Cleaning	o 4 cleaners clean the building on each morning before the Café opens
Management	o 1 manager (full-time) o 2 assistant managers (full-time). They do the planning, purchasing, personnel administration and handle legal issues. o 1 assistant manager (part-time). For the administration, including the financial administration.

Table 1.2: Departments and jobs in a Grand Café

With this short piece of information, we invite you to different challenges for understanding the diversity of information and ways of managing information and how this helps to solve actual business problems in the private or public domain, profit or nonprofit. See Table 1.3 for some key questions that are related to each inquiring systems and assignment.

Inquiring system	**Key questions**	**Related sub discipline and business issues**
A1. The Lockean inquiring system and databases	What data do managers need for their business (e.g. on prices, suppliers, performance of products, numbers served in different places and regions)? What are the relations between these data? What reports are most useful for whom? How can we use software to maintain data (consistency) and make management reporting easier?	Human resource management; marketing databases; product and inventory management; project team sourcing; quality management
A2. The Leibnizian inquiring system and decision models	What are the causal relations between decision parameters? What parameters are important to predict and forecast reality (e.g. number of expected clients; availability of waiters; specialisms, qualifications and experiences of kooks and waiters?). How to solve long waiting times?	Production planning; procurement decision making; decision science; logistics; budgeting; accounting
A3. The Kantian inquiring system and multiple perspectives	What are the most relevant category and method for modeling a business process or organizational structure? What processes jointly make up the production or service processes of an organization? How are business processes related and integrated (if at all)? Who are involved of each process? What is the input and output of what processes? How are tasks divided in the department? Is it possible to make processes more efficient, lean, and possibly automated?	Logistics; process design; operational management; organization design; project management and planning
A4. The Hegelian inquiring system	How is information used in organization-political contexts? How can we detect and	Law; corporate communication; performance appraisal;

and information politics	correct information biases? How should one correctly interpret indicators (e.g. a waiting time score)? Is it correct what is said by the media and the market about us or others?	management by objectives; marketing communication; innovation management; new product development and improvement; quality management.
A5. The Singerian inquiring system and problem solving	How can one collect information from multiple sources that can be made useful for problem solving, product and service improvement?	Marketing; Product and service innovation management, creativity and problem solving.
A6. The organizational context of information	What departments and persons need what information system? What are the relation among these systems? What is a useful information plan and strategy for the company? How do the information systems allow a firm to achieve its strategy? What systems make an organization vulnerable for process breakdowns? Do the existing systems well fit with organizational needs and capabilities?	Organization strategy; organization capabilities and competencies; organizational change; project management and planning; organization design; new organizational design parameters.

Table 1.3: Information concepts, assignments and QuickStart

Although we touch on fundamental, and even philosophical, concepts and theories, the text should be evaluated on its practicality in developing entrance level skills for students of management and administration. This text is especially suitable for entrance level students in business, organization science, public administration, economics and social sciences, but could serve anyone with a practical insight in what modern IT could mean for them. Table 1.3 for example gives in its third column a set of example links with different other management sub-disciplines.

We will illustrate most of the related software skills by Microsoft products: MS-Access, MS-Excel and MS-Visio. This is neither because we regard Microsoft products to be superior to others nor that we have any commercial interests in their products, but simply because they are most widely distributed and have

been trendsetting in functionalities and user-interfaces for many other business software products. Most universities also have special financial arrangements with the Microsoft Corporation by which their software is available for very low rates. For Apple machine users, Microsoft offers more expensive variants of their software, although one may also run the Windows operating systems parallel to Apple's OS by which cheap MS products become available for them. Most modern Apple machines allow running parallel operating systems without much delay.

This book is in English. If you have Microsoft products in another language, note that you can easily switch the language as follows. Select under windows 7: Start, programs, Microsoft Office, Microsoft Office Tools, Microsoft Office Language preferences. You will see the following screen:

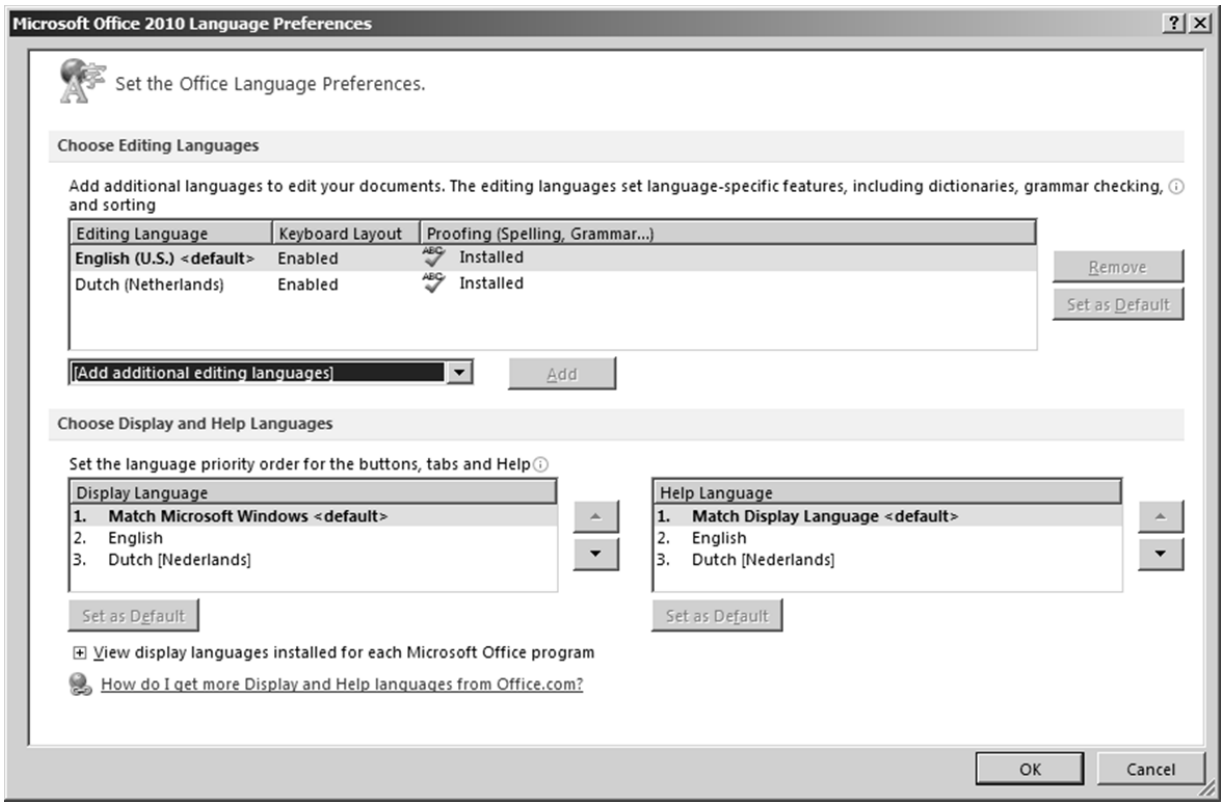


Figure 1.2: The Microsoft Office language selection screen

Here, you can choose the default languages as you wish. In the screen above, the default language is English.

I do not believe that all information in this book is complete, perfect and up-to-date (regarding the last this is impossible given the fast changes in the IT industry), and it gives certainly not the last word about information management. But I do believe that it gives a skeleton of thought regarding information and its management that will be of value to students during the rest of their career. This is especially achieved by linking information concepts and methods to fundamental philosophical thought. Therefore, each information skills will be in each chapter introduced by some key points of the related inquiring system.

I hope that the readers enjoy reading and working with this book. The experiences of the readers with it will be used to make an enhanced version in the future. For this the readers and especially the users of this text are much invited to send suggestion for improvement to fons.wijnhoven@utwente.nl.

2. Empiricism, data management and databases

2.1 Locke and databases

Empiricism (based on the philosophy of Locke) is a theory of knowledge emphasizing the role of experience, especially sensory perception, in the formation of ideas, while discounting the notion of innate ideas. The Lockean inquiring system is based on the philosophy of John Locke (1632-1704) (Uzgalis, 2010), who postulated in his first book on "An Essay Concerning Human Understanding" that there are no innate ideas that form our understanding of the world. In book II, he states that the mind is a "blank slate" or "tabula rasa"; that is, contrary to Cartesian or Christian philosophy, Locke maintained that people are born without innate ideas. In book III, he states that language is a key element in forming and codifying understandings, and that we need to share common meanings to make knowledge sharing feasible. Locke recognizes that ordinary people are the chief makers of language and that scientists have the task of checking if the connections made between properties in reality in this language are actually true or not. In book IV Locke states that man should try to use reason, i.e. a combination of observation, experience and rationality in finding truth. But, people have limitations in reasoning, because many issues are too complex (for the time being), and in such cases it is reasonable to believe. So revelation comes in where reason cannot reach (Uzgalis, 2010). According to this inquiry system, knowledge is an intersubjective representation of the world and believed to be "objectively" true by consensus in the community of its owners. Information failure in this context results in a lack of correspondence with reality.

Empiricism (Locke's approach) emphasizes true facts about reality as the key to understanding the world. Our senses are a data collection mechanism, and the data we collect should be more or less directly understandable and be shared with other people to develop a collective understanding. Data collection and knowledge creation, though, require substantial work of the mind. The mind for instance will have to develop interests in specific areas (trying to collect all data about everything is obviously impossible) related to specific challenges and goals (Berger & Luckmann, 1967). These interests and goals thus delimit the boundaries of our universe of discourse, which in turn specifies the data needed and the mechanisms by which these data can be efficiently and effectively acquired (Halpin & Nijssen, 1995). These data need interpretation and analysis to create relevant information for decision making and problem solving, but an important bias is already created by the sensory mechanism itself (Berger & Luckmann, 1967). A universe of discourse description contains meaningful coding and words (so called

semantics) for representing relevant phenomena. The Lockean inquiring system thus focuses on the representational veracity, completeness and meaningfulness of primary data, i.e. data that are direct representations of reality.

According to the Lockean inquiring system at least four steps are required to make data meaningful (see Figure 2.1):

1. Scoping and identifying the objects in the world about which data have to be registered.
2. Data definitions as the descriptions of labels of objects and for standardization of the meaning of data in a database.
3. Data models for an efficient way to deal with data, preserve consistency among data and maintain the data;
4. Reporting mechanisms that allow to (quickly) produce relevant reports from (large amounts of) data.

Data definitions and data models thus are data about primary data, and therefore also named meta data.

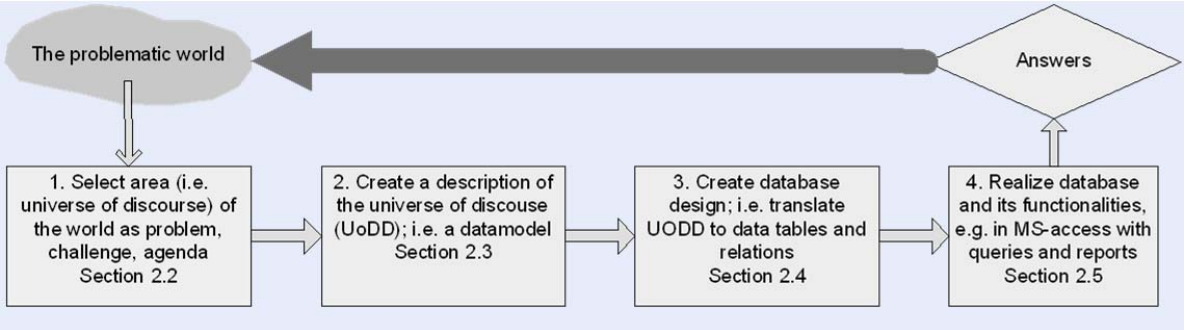


Figure 2.1: Steps of the empirical approach for making data meaningful

A database is an explicitly logically structured collection of data. When using MS Access, you can relatively easily make a database. However, first it is important to design the database in a number of steps. Note that following the four steps mentioned above, the relation between people's observations and sensations and the actual creation of data which may be managed by a database is intermediated by sensory mechanisms and labels for real events. This is also explained by how initial observations are collected and transformed to data that astronomers can analyze and interpret as a "gamma ray burst" (see the Wikipedia page on this topic on http://en.wikipedia.org/wiki/Gamma-ray_burst, accessed November 27, 2013). The gamma ray burst is an "entity" that we perceive to exist in reality, and which has certain labels (named "attributes" in the database discipline) like location, a certain strength, length and duration depending on the scope of the (scientific) language we apply to describe and analyze the phenomenon.

SWIFT is the telescope, thus the sensory mechanism, which sends out unlabeled sensations via satellites to ground stations, where the signals are next labeled and its bundle of attributes is induced as an event (entity) named "gamma ray burst" and connected with time, location, duration and other attributes to make an analysis of what actually happened. These analyses also can be used by scientists to create knowledge on:

- The origin of gamma-ray bursts.
- A classification of gamma-ray bursts and search for new types.
- Determining how the burst evolves and interacts with the surroundings.
- Using gamma-ray bursts to study the early universe.
- Performing the first sensitive hard X-ray survey of the sky.

Note that the identification of a set of observations as "gamma ray burst" is impossible without a language that has a word and definition for this entity. This language is of course not independent of the development of our theories of the phenomenon and what we want to include or exclude from this phenomenon (of course for astronomers, not all phenomena are gamma ray bursts). The whole set of entities that people are interested in and the labels/attributes they use to talk about them is named the "universe of discourse" (Berger & Luckmann, 1967). Universe of discourses are mostly related to our everyday use of language (Anderson & Prelli, 2001). When we aim at using computers to help us with registrations, measurements and analysis (and we often need this for achieving efficiency, speed and quality in decision making and analysis), we need to be able to represent the universe of discourse in an unambiguous way and such that computers can automatically process their "metadata", because the meta-data define the meaning, also named semantics, of the primary data. If people give computers the responsibility to manage these universe of discourse descriptions, they have to be formal, precise and disambiguous.

A similar mediated situation between organizational events and understanding of them exists. For example, a purchase has many necessarily related entities to understanding it, like the moment and place of purchase of the product or service bought, a price, a volume, a sales person, a buyer, and a payment. All these entities need attributes to register them. A sales person can be registered via his or her name and employee number. A price must be expressed in a number that represents units of money of some type. A buyer can be registered by an account number, name, and address. Similarly, social media enable the identification of unique events or characteristics (profile) of a person by so-called EXIF labels to the photo's people upload on for example Facebook. This EXIF thus includes GPS coordinates (with 15 ft

precision), and data and time information that both are automatically attached to uploaded photos, unless you install specific privacy settings on your phone.

2.2 Goal of this chapter

This chapter will explain how to make a Universe of Discourse Description (UODD), which enables you to manage large volumes of data consistently via database software for the generation of relevant (management) reports. To realize this we discuss the following:

- The Universe of discourse (UoD) as a set of related entities about which you want data in the database (see section 2.3).
- The Universe of discourse description (UoDD) as a set of entities with attributes being the meta-data and definitions of relations among the data (see section 2.4).
- The Database design by MS-Visio (see section 2.5).
- Implementing the Database design in MS Access (see section 2.6)
- Making of a report by MS-Access (see section 2.7)

Note that if the UoD or UoDD are incomplete, it will be impossible to make complete reports. However, if the UoDD includes too much, it will become very expensive and complicated to maintain its quality and consistency, and no one is happy with too much irrelevant data, i.e., information overload (Landau, 1969).

The universe of discourse is that part of the world that is central to the problem for which the database is needed. A problem is for example the need of a Project Manager to have overviews of who is working on what tasks and to see how much budget and time is still available for the project. A relevant UoD thus may include “employees”, their “hour rates”, “tasks” and “departments”. The information need of the manager consists of relevant reports that can be produced from data about these entities. The universe of discourse determines the scope of the problem and consists of a number of entities (and excludes a lot more entities of the world). These entities are the core subjects over which data is collected.

The information objects may be entities, their attributes, or their values. These have the following characteristics…

1. Entities are phenomena of reality, like persons, cars, and countries. The whole set of relevant entities and their logical relations is what we call the Universe of Discourse.
2. Attributes are characteristics of entities. The entity “person” may have attributes like “name” and “date of birth”.

3. Entities and attributes have logical relations. For example entity “China” has an attribute “population”.
4. Attributes have values. For example the value for “population” for China is “1.3 billion”. Some attributes are unique for one entity and thus can identify it. Such an attribute is named an “identifier”. Population number is not a good candidate for identifying a country. Although not many countries have a population of 1.3 billion people, some countries may have the same number of people. For identifying countries uniquely they need to have a unique name or code that is accepted and used by all users of the information system. For example a person’s passport number such identify a person in a unique way and should not be used by any other person in the world to identify him or her.
5. Entities can have relations as well. For example China and the Netherlands “collaborate on the exchange of students”.
6. Attributes can have meaningful relations. For example sometimes multiple attributes identify an order in a unique way, e.g. a client code, a product name, and a date. Often we want reports of entities and multiple attributes, for example students with names and the classes they join.
7. Values can have correlations. For example a higher income may correlate with a better health condition or if your income is below X, you may apply for a study allowance larger than Y.

The set of entities, attributes and logical relations is what we call the Universe of Discourse Description (UoDD). In the database literature UoDD are also named Entity Relation models, which is a more general term for metadata models. These meta models are themselves not necessarily based on the empiricist philosophy, because they are logical and mostly based on set theory and mathematical theory (Hoffer, Prescott, & McFadden, 2007) or linguistics (Halpin & Nijssen, 1995) which we will not discuss here, but they are preconditions for the purposeful and organized management of observations. Correlations among values of attributes, however, are an important type of scientific result aimed at by empiricists, besides of blunt measurements and data.

2.3 Making a UoD using MS Visio

MS-Visio offers some very useful tools for expressing a UoD. For a UoD, one only describes the entities and the relations between them. So, attributes are not mentioned in a UoD. Using MS-Visio, this is possible in the following steps…..

- Open MS Visio
- MS-Visio has a lot of templates, which are sets of related shapes for making any kind of model. For example, there are templates for floor maps (with shapes like doors, desks, and chairs) for organization charts, and for business process maps. The template, mostly named a “stencil” by MS-

Visio, used here for UoD's is the Crow's Foot Diagram. You can find it at 'Software and Database' see Figure 2.2.

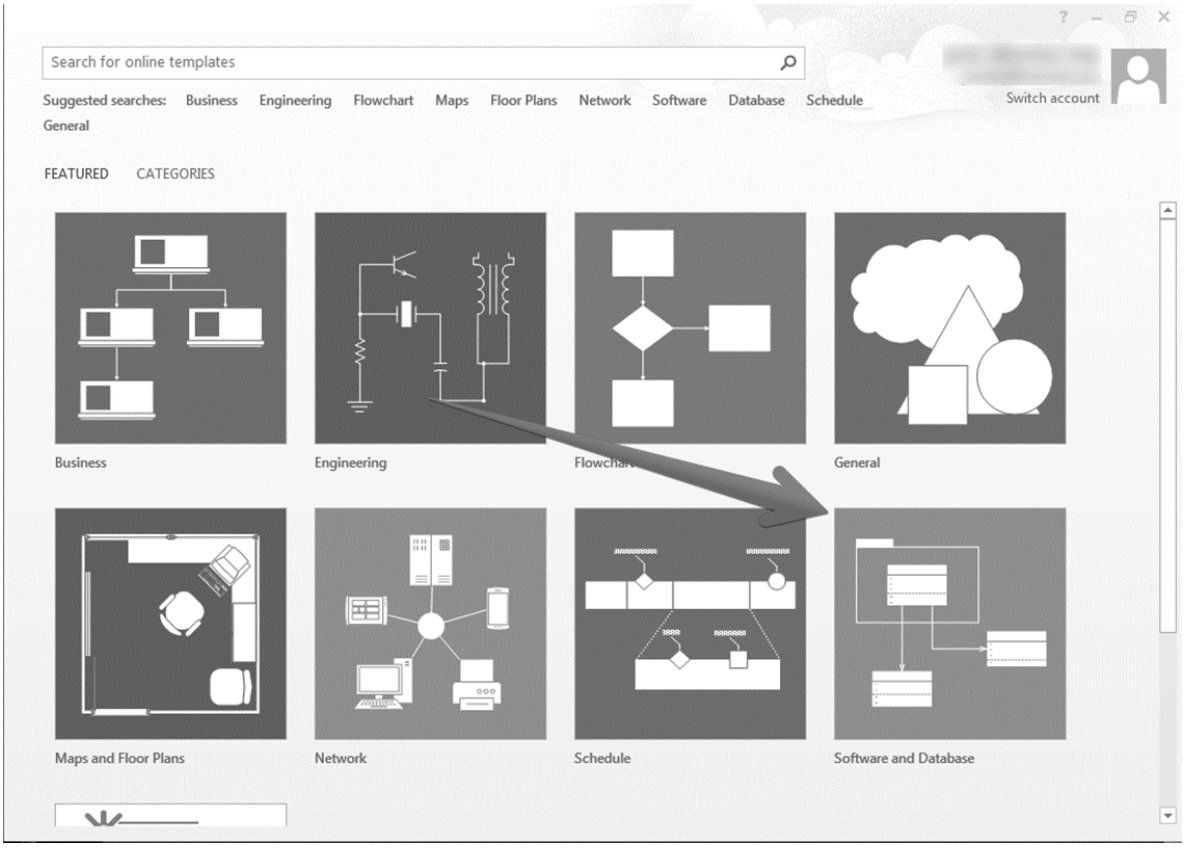


Figure 2.2 Software and Database menu in Visio 2013

- When you open the Crows Foot stencil, a worksheet/drawing page appears. When you want an entity-shape on your worksheet, you have to drag an entity-shape from the left bar to the right side of the screen (your worksheet). While the shape is still selected, you can change the name of it (e.g. employee). By convention, all entities have a singular name (e.g. "employee" instead of "employees").
- When all entities are on the drawing page, it is time to connect them. For doing so, one need to know which entities are logically and meaningfully related according to your view of the world. Connect the entities with a 'connector' (in this case a line without an arrow at the end). When you have selected the connector tool from the "Home" tab, hoover the cursor over the entity. A green square will appear to indicate that you can connect the connector here. Now keep pressing the mouse button and hoover over the other entity and make the connection. The entity now can move, without

breaking the connection. Now connect the other end from the connector with another entity. After all entities are logically connected, the first step of database design is completed.

See an example of a possible UoD for a project manager in Figure 2.3.

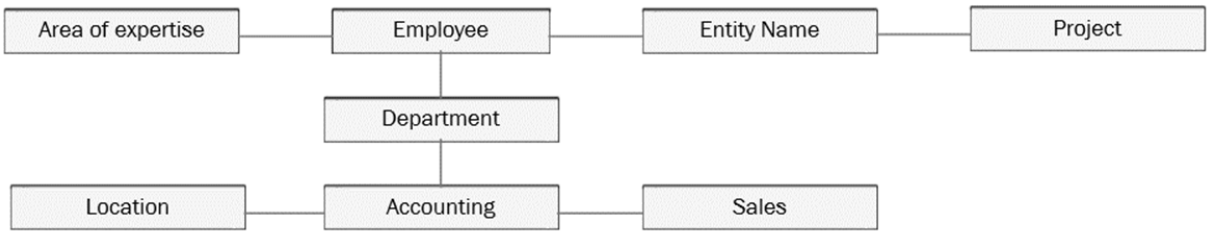


Figure 2.3 Example of a UoD.

2.4 Universe of Discourse Description (UoDD)

A universe of discourse description (UoDD) is a specification (i.e., a precise disambiguous description) of the UoD. In a UoDD, attributes are connected with the entities, which describe the type of information that will be registered in the database. Attributes describe the characteristics of an entity. For example, a student (entity) in a student registration database may have the following attributes: “unique student number”, a “name”, a “telephone number”, an “address”, “a study program” in which s/he may be enrolled, and an “e-mail address”. The unique student number is called the identifier (primary key) in database language. When one types this number in the database, only the record of this student will be retrieved, because s/he is hopefully the only student with that number.

In business, for example, client numbers, order numbers, supplier number, and part number have to be unique. (Just imagine what would happen if it would not be the case). Furthermore relations exist between an entity and corresponding attributes and among entities. Using the Crow’s foot diagram, we can create these relations as follows:

- Open the UoD diagram (mentioned above) in MS-Visio.
- Click in the MS-Visio menu bar on ‘insert’ and ‘new page’. Name this page (e.g. UoDD). Now you have a new drawing page, but you still work in the same map.
- Copy the entities from the UoD on this worksheet. First you have to connect the entities in the same way as you did in the UoD. But, the connectors now will have another shape, the original crow’s foot shape, because you also have to describe the logical relation. These relations are called cardinalities.

The cardinalities are expressed by the way the connector connects to each entity. As you can see in Figures 2.4-2.7, this looks like a crow's foot.

	This figure represents a "one to one" relationship between the entities. This is indicated by the two vertical stripes, which means it can be "one and only one". Example: the relation between a person and a social security number. Every person can have only one social security number and the other way around: every social security number can belong to only one person.

Figure 2.4: one-to-one relation between entities

	In this figure the symbol on the right indicates a "one or more" relationship. This means that the entity on the left always is one, while the entity on the right can be one or many. Example: the relation between a department and an employee. Every department can have multiple employees, but every employee can be member of only one department.

Figure 2.5: one or more relation

	In this figure the symbol on the right indicates a "zero or one" relationship. Example: the relationship between a student and a registration for a subject. Every student can have no or one registration for a specific subject.

Figure 2.6: zero or one relation

	The symbol on the right in this figure indicates a "zero or more relationship". Example: the relation between a student and the classes he takes. Every student can take zero or more classes during his study period.

Figure 2.7: zero or many relation

Besides the relationships, you can also give attributes to each of the entities. Drag the attribute item from the left sidebar onto an entity and it gives you the possibility to define attributes for this entity. As explained before, it is also important to think about assigning Primary Key's to identifier attributes, which make sure that every entity is unique. See figure 2.8 for an example of this.

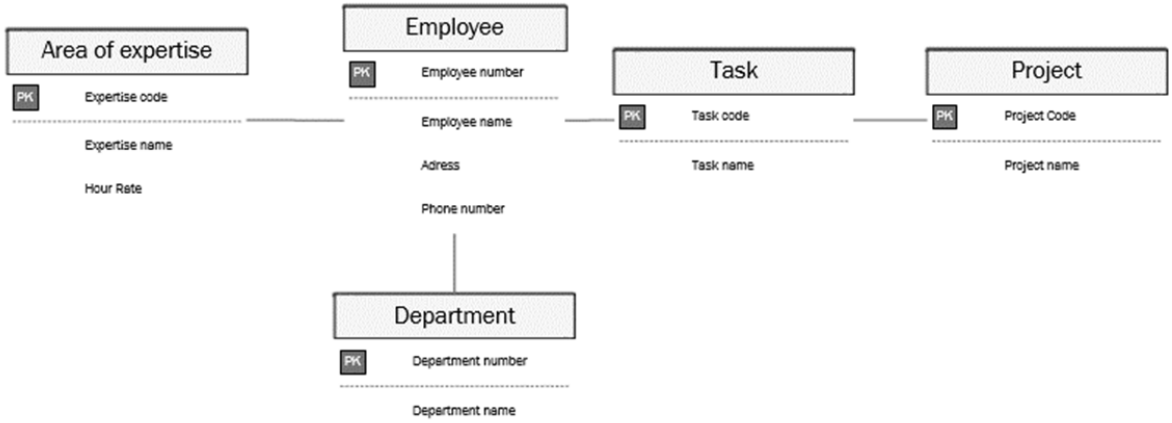


Figure 2.8 Universe of Discourse Description

Now you have a full Universe of Discourse Description. The UoDD must meet the following requirements:

- Use correct shapes for entities and attributes;
- Consistent with UoD: use of the same entities and relations;
- Attributes are defined;
- The relations among the entities and the entities and attributes are "logical" and made with the correct method.

2.5 Implementing the database design in MS Access

Now, we can implement the database design in a database management software package (like MS Access):

- Open MS-Access.
- Open an empty database and name it (e.g. database Personnel Manager). We do not use a standard template in MS-Access because we already made our own database design.
- Create Tables. Every Table from the database design needs also a Table in MS-Access. In MS-Access 2007 you will have the screen of Figure 2.9.

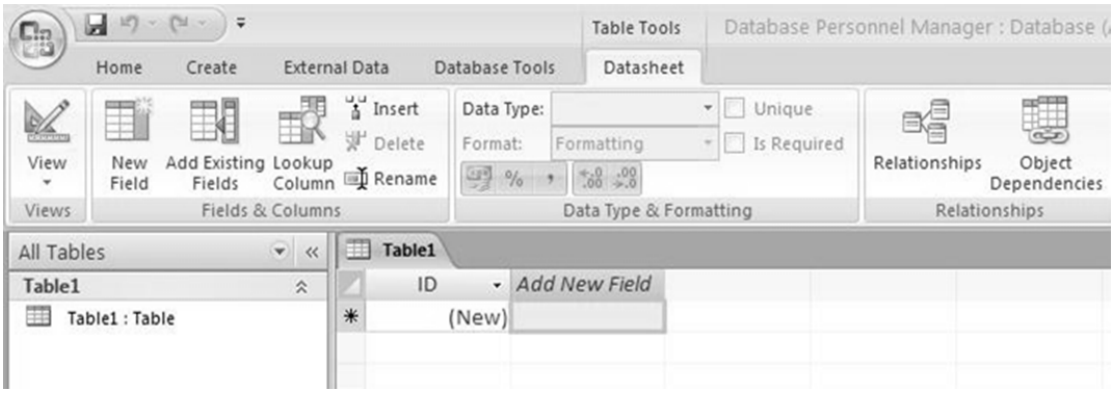


Figure 2.9: The Table creation screen in MS-Access

Click on View and Design View (see Figure 2.10).

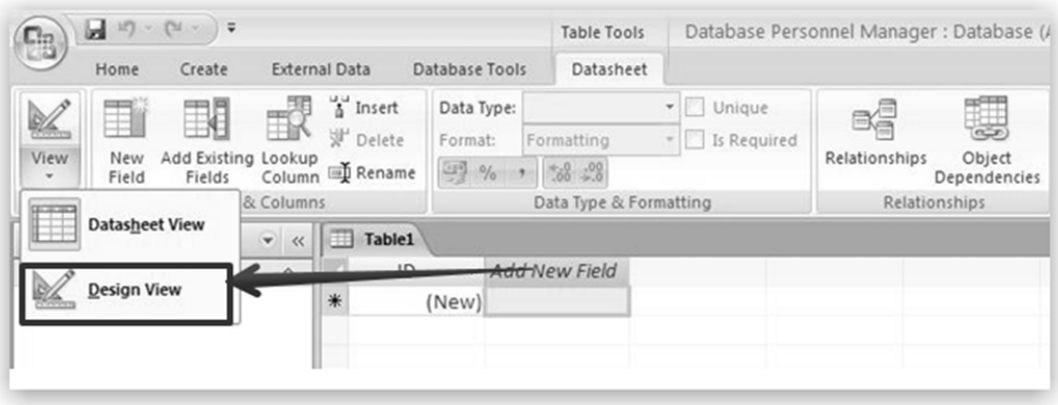


Figure 2.10: Selecting the data design view in MS-Access

You will have the following message of Figure 2.11.

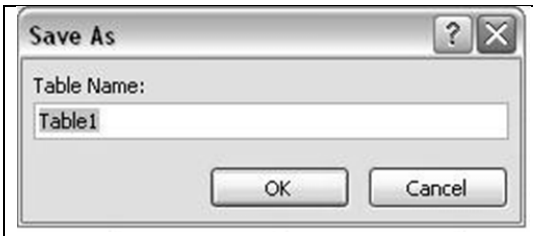


Figure 211: Table creation

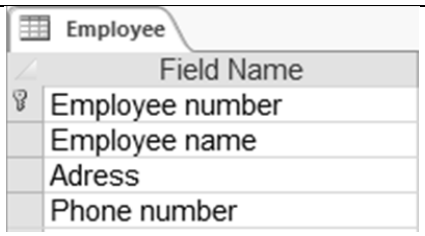


Figure 2.12: A Table in MS-Access

- Name the Table, e.g. Employee, and click on OK.
- Type (under “field name”) all attributes, beginning with the primary key (Figure 2.12). In our case this is an “employee number”. Also type all other attributes in the MS-Access field. These attributes have to match with the concerning table from the database design made in MS-Visio. So you have to type the foreign keys as well! The employee table now looks like in Figure 2.14 in the database design.

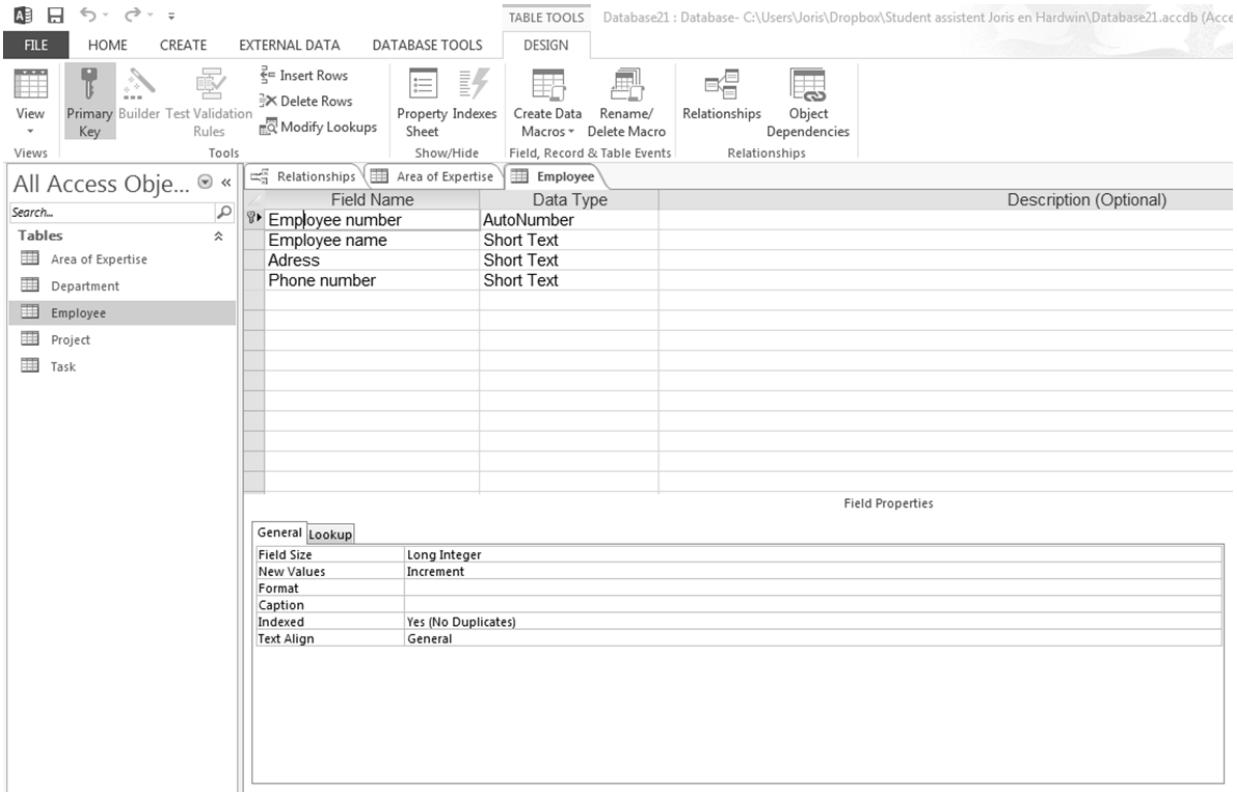


Figure 2.13: Allocation of the primary key to a field name in MS Access

In Table 2.1 you can read the explanation MS-Access gives about the second column of the table: data type. The attribute Data Type can have the following settings:

Data Type	Used To Store	Limitations/Restrictions
Text	Alphanumeric data (text & numbers)	Stores up to 255 characters.
Memo	Alphanumeric data (text & numbers)	Note that adding 2GB of data causes your database to operate slowly. If you enter data manually, you can enter and view a maximum of 65,535 characters in the table field and in any controls that you bind to the field. When you create databases in the Office Access 2007 file format, Memo fields also support rich-text editing.
Number	Numeric data	Number fields use a Field Size setting that controls the size of the value that the field can contain. You can set the field size to 1, 2, 4, 8, or 16 bytes.
Date/Time	Dates and times	Access stores all dates as 8-byte double-precision integers.
Currency	Monetary data	Stores data as 8-byte numbers with precision to four decimal places. Use this data type to store financial data and when you do not want Access to round values.
AutoNumber	Unique values created by Access when creating a new record	Stores data as 4-byte values; typically used in primary keys.
Yes/No	Boolean (true or false) data.	Access uses -1 for all Yes values and 0 for all No values.
OLE Object	Images, documents, graphs, and other objects from	Stores up to 2GB of data (the size limit for all Access databases). Adding 2GB of data causes your database to operate slowly. OLE Object fields create bitmap images of the original document or other object, and then display that bitmap in the table fields and

	Office and Windows-based programs	form or report controls in your database. For Access to render those images, you must have an OLE server (a program that supports that file type) registered on the computer that runs your database. As a rule, you should use Attachment fields for your .accdb files instead of OLE Object fields.
Hyperlink	Web addresses	Stores up to 1 gigabyte of data. You can store links to Web sites, sites or files on an intranet or Local Area Network (LAN), and sites or files on your computer.
Attachment	Any supported type of file	You can attach images, spreadsheet files, documents, charts, and other types of supported files to the records in your database, much like you attach files to e-mail messages. You can also view and edit attached files, depending on how the database designer sets up the Attachment field. Attachment fields provide greater flexibility.

Table 2.1: Data types in MS Access, retrieved from (http://office.microsoft.com/en-001/access-help/modify-or-change-the-data-type-setting-for-a-field-HA010096450.aspx)

We use AutoNumber to give identifier values, because this is a unique sequential number (that increases in steps of 1) or a unique random number, which is allocated to every new record which is added to a table. Fields from the type AutoNumber cannot be updated. For other attributes from employee itself we use 'Text'.

For foreign keys we do not use Text but Number values ("Numeriek"), because this has to be equal to the settings in the other tables (otherwise there will be problems when you have to make connections). 'Text' means that the field automatically has a size from 255 characters. For something like a phone number you only need 10 characters. It is easy to change the field size to 10, this will also prevent you from filling in a too long phone number for example. Enter at the bottom 10 at field size. You can do the same with for example a zip code.

A next step is to give descriptions to all attributes (see Figure 2.14).

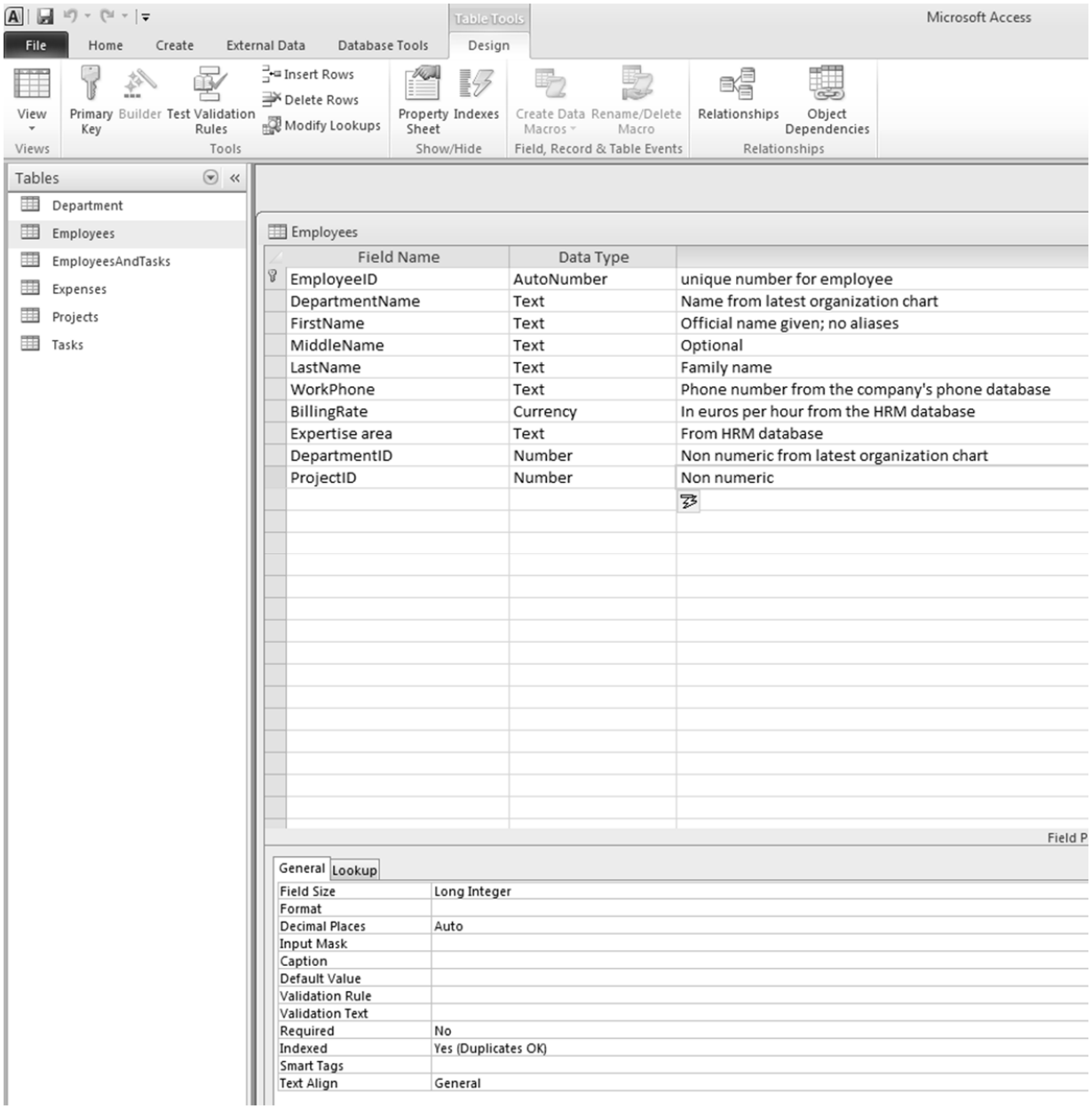


Figure 2.14: Giving descriptions to attributes.

Using the steps as described before, create tables for all 5 entities (area of expertise, department, employee, project and task). The next step is creating connections among the tables:

1. Save all the tables at the background (right mouse click on the tab of the table and ‘Save’)

2. Close all tables at the background (right mouse click on the tab and ‘Close’). Otherwise we cannot make connections.
3. Go to ‘Database tools’ and choose ‘Relationships’ (see Figure 2.15).

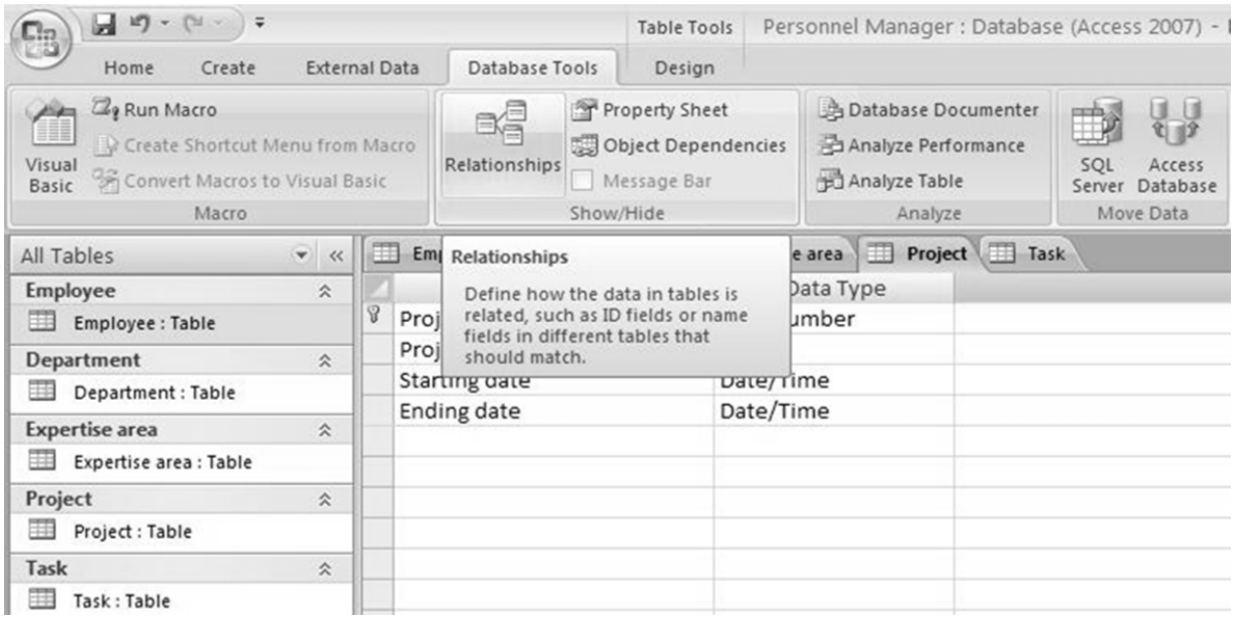


Figure 2.15: Select the relations tool in MS-Access

4. Add the tables by selecting them and click on ‘Add’. They pop up in the back of the screen.
5. When tables are added, click on ‘Close’.

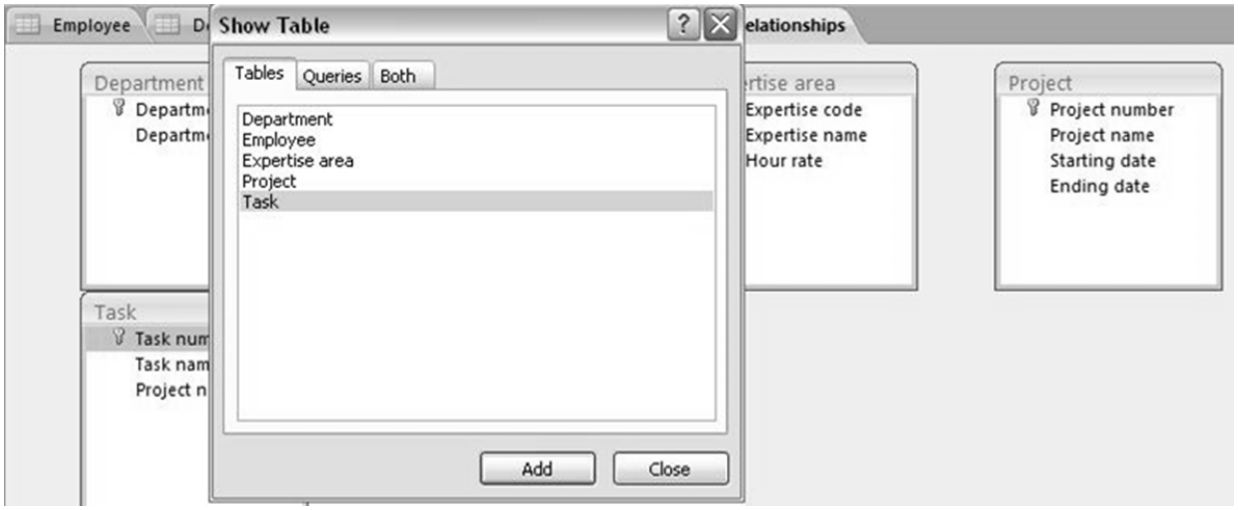


Figure 2.16: See what Tables can be related

Now we are in the screen with only the tab ‘Relations’, and the five tables. Now we have to implement the connections between the tables in four steps:

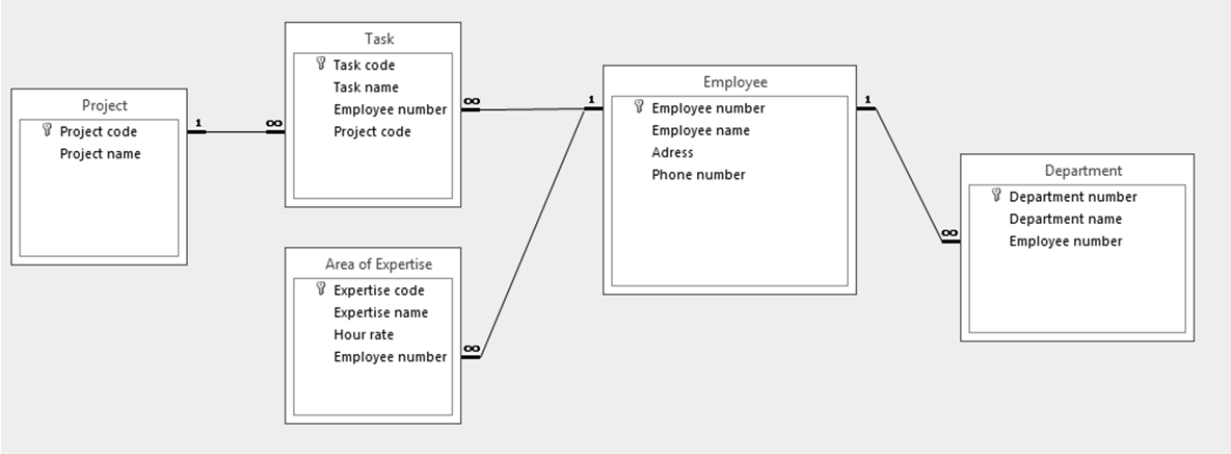


Figure 2.17: Table and their keys fields, including the relations and cardinalities

1. Drag the primary key from the one table on top of the foreign key from the other table. Click for example on ‘department number’ from the table ‘department’ and keep the mouse pressed. Go to ‘department number’ from the table ‘employee’ and release.

2. In the screen you get, you have to check ‘Enforce Referential Integrity’. Check also the other two boxes. The screen now looks like Figure 2.18.

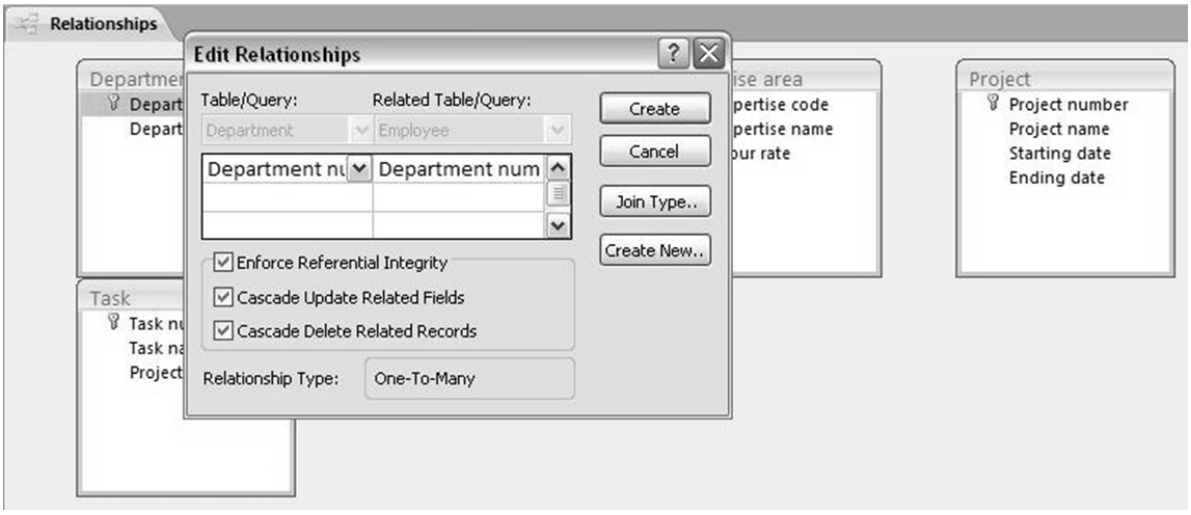


Figure 2.18: Edit relations and add cardinalities

3. Next, click on ‘Create’. Now there is a line between the two tables with on the one side a ‘1’ and on the other side the infinite character “∞”. This means a 1:N relation.

4. Repeat this for all the tables. In our example it looks like figure 2.17.

IMPORTANT: In the tables you have just built, all the relationships are 1:N. However, since an employee can have multiple tasks, and multiple employees can work on one task we need an M:N relationship. This is also the case for all the other relationships. Unfortunately MS-Access only allows a M:N relation by inserting a third table. This is the so-called connection table. Such a table contains the two primary-key fields from the two connected tables. For example the connection table connecting an employee to a task contains to columns, the primary key for task code (from the table Task) and the primary key employee number (from the table Employee). See Figure 2.19.

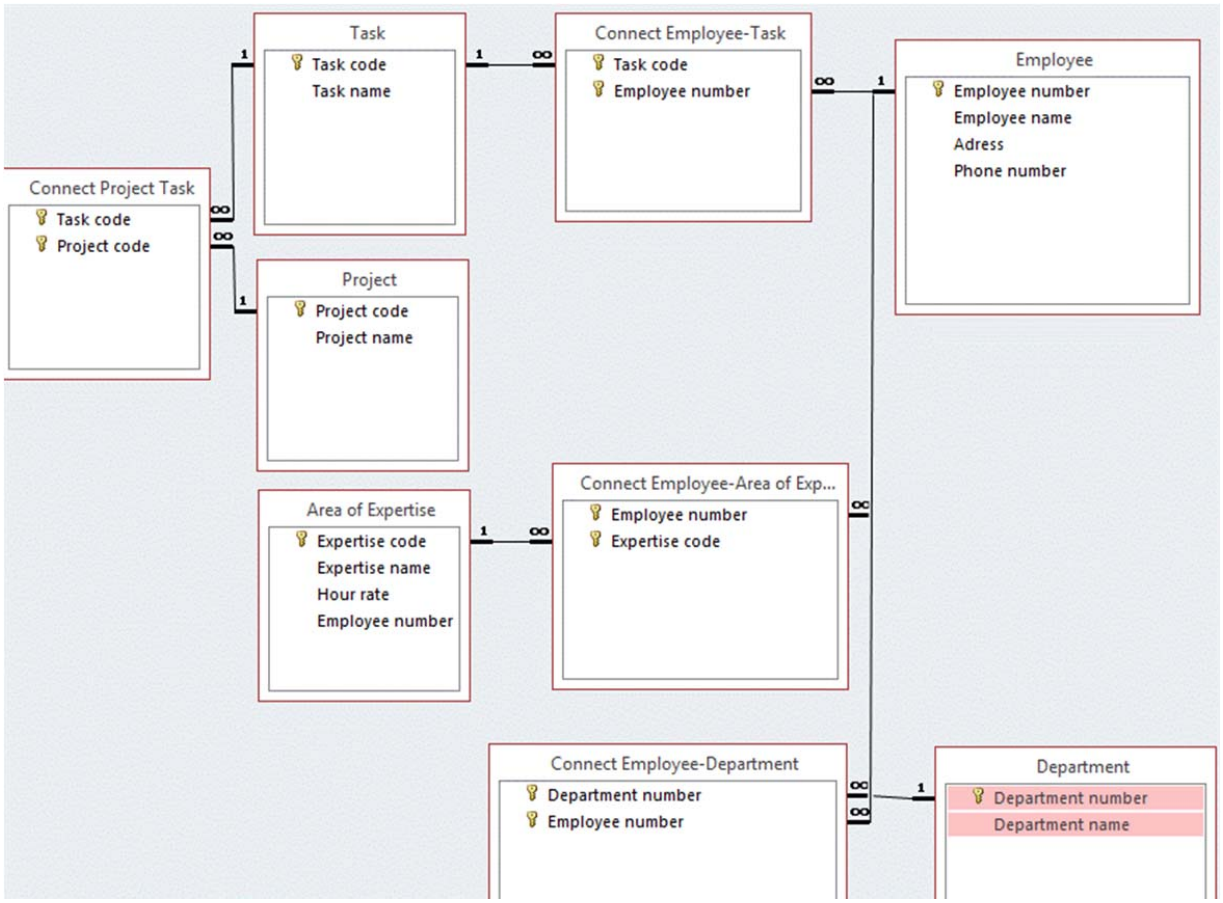


Figure 2.19 Database design with relations and connection tables

When all the connections are made, you have to fill the tables with data values as in Figure 2.20. First we fill all tables that do not have FK's. These are department, project and expertise area. Next we fill out the "task" and after that the "employee" table. This is because MS-Access immediately makes links to other tables. If underlying data is not yet created (for example in the case of a department number), MS-Access cannot make these connections. For this:

1. Double click on a table left under 'All tables' (so first the department, project or the expertise area table).

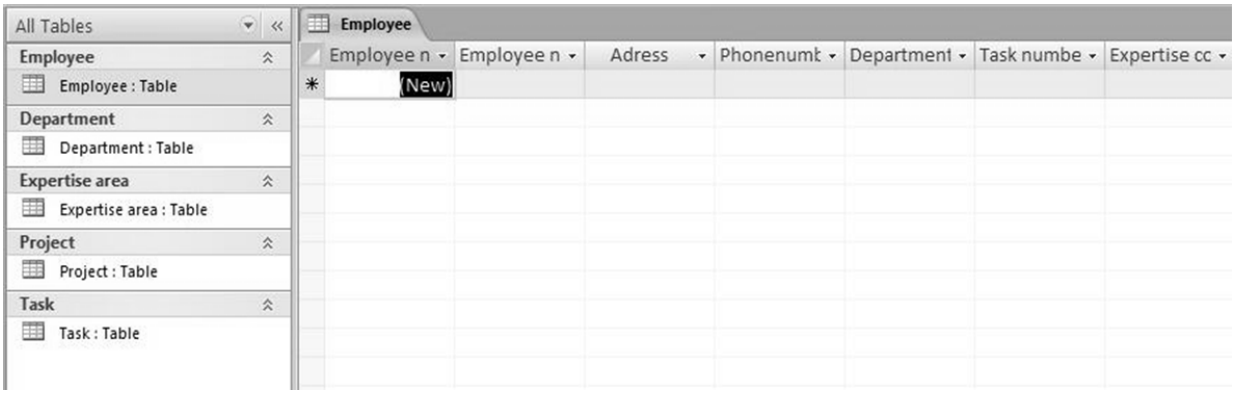


Figure 2.20: Opening a Table for data entry

2. Enter data in the input screen.

At AutoNumber we do not have to enter anything; the numbering will be done automatically (therefore 'auto' number). Further, for example, we insert at the table department the department names. We do this just for all the fields that we made.

It is much work to maintain consistency. For example, you cannot fill in '100' at task number in the employee table, because that means you have to define at least 100 tasks (because the AutoNumber starts at '1'). It is your own choice how much data you insert, but to make a useful report, you have to insert at least 5 records everywhere. Here are some examples for two tables.

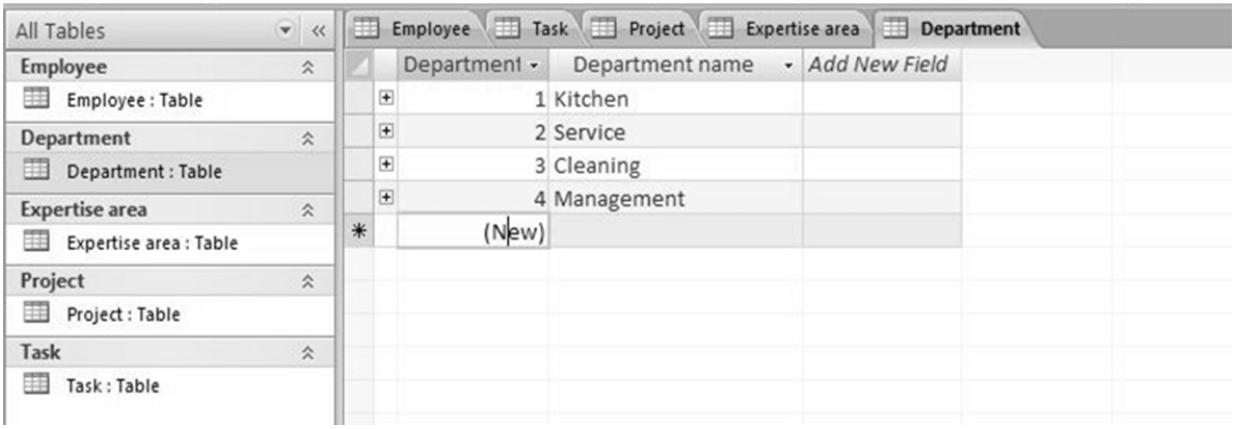


Figure 2.21: Data entry Table "Department"

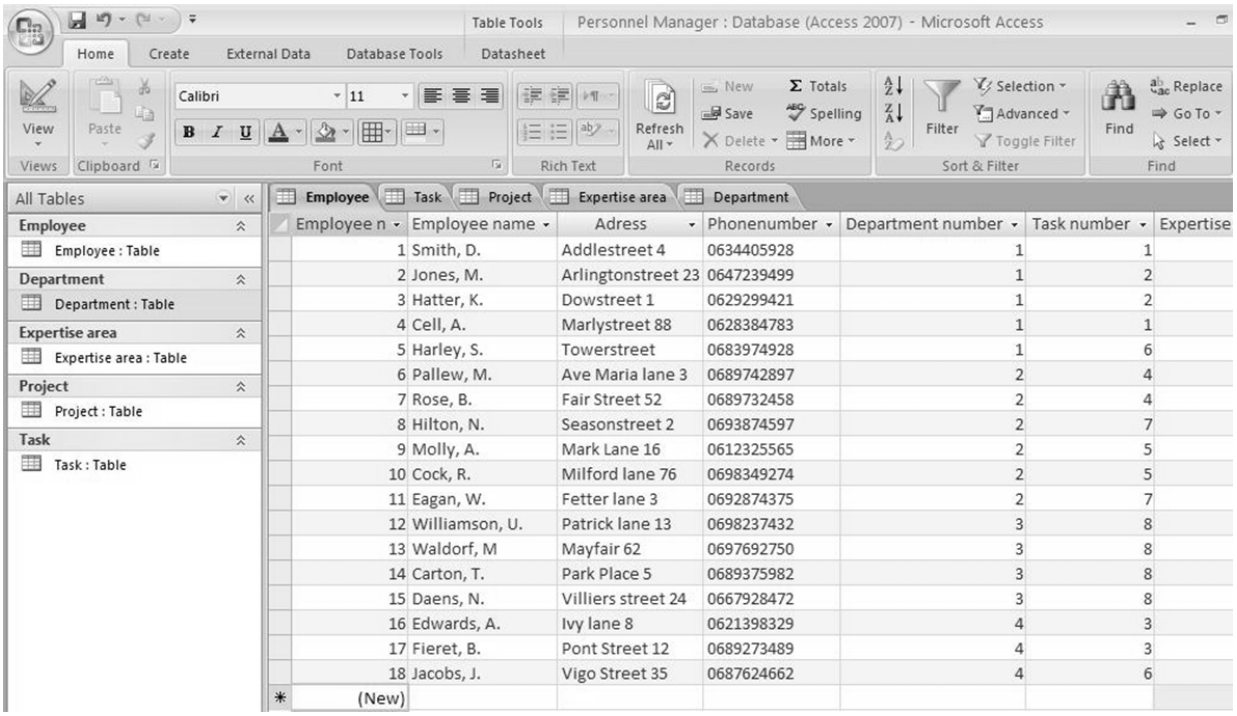


Figure 2.22: Data Entry for table “employee”.

3. Save everything.

2.7 Making an MS-Access query and report

We will start with making a query, and continue with making a report after that in this section.

Click in the menu bar on ‘Create’ and ‘Query Wizard’ as in Figure 2.23.

Figure 2.23: The query wizard

Insert ‘fields’ (attributes) (A) from different Tables (B), due to the attributes you need for your report.

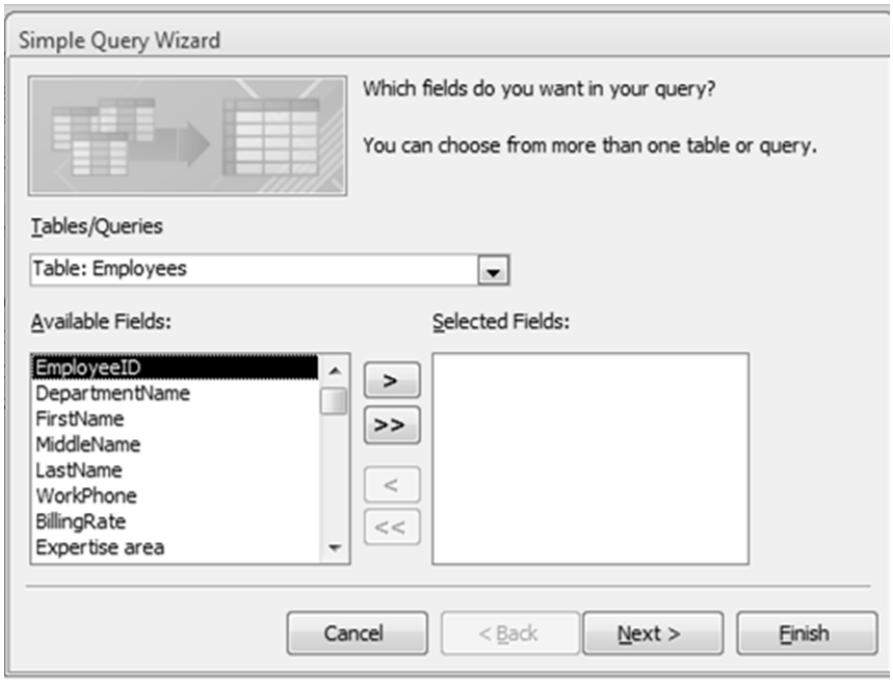


Figure 2.24: Selecting field for the query

Suppose the Personnel Manager has a question: which employees are performing what tasks? He has an Employee Table, a Task Table, and also knows that Employees and Tasks have a 1:1 match.

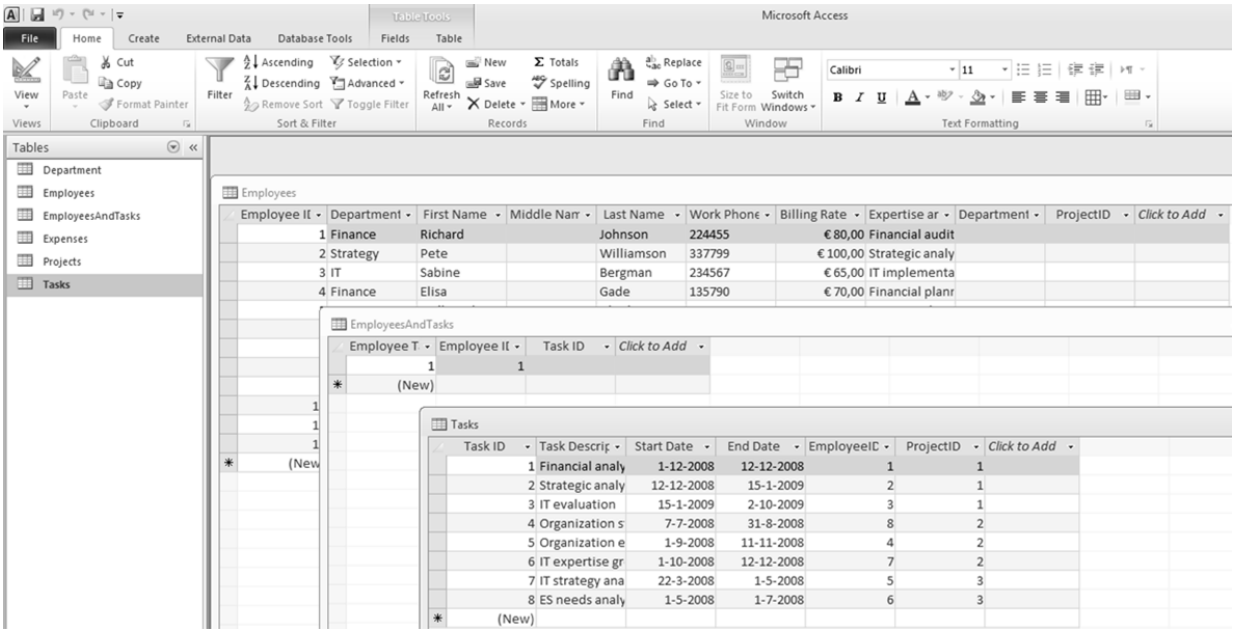


Figure 2.25: The MS-Access Tables from which a report can be generated.

To develop a query from the Employee and Task table, Fields are selected from both tables to be included in the query.

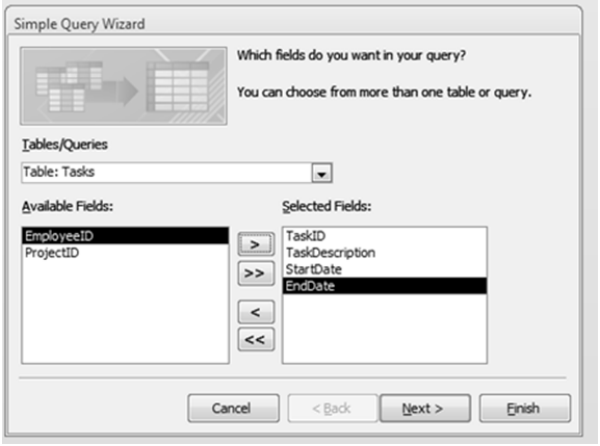


Fig. 2.26: Fields from the Tasks Table

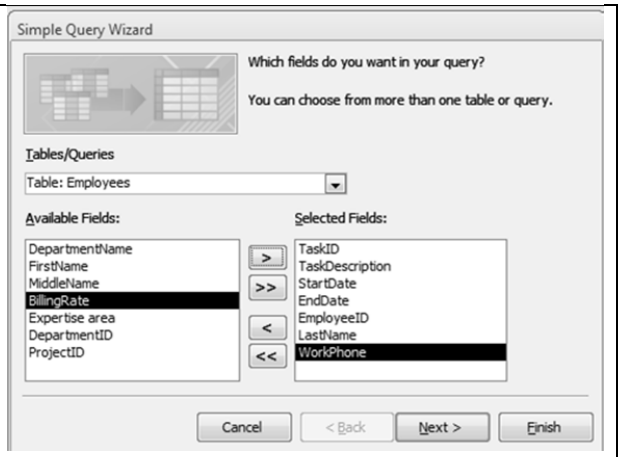


Fig. 2.27: Fields from the Employees Table

After pressing “finish” the following screen pups up.

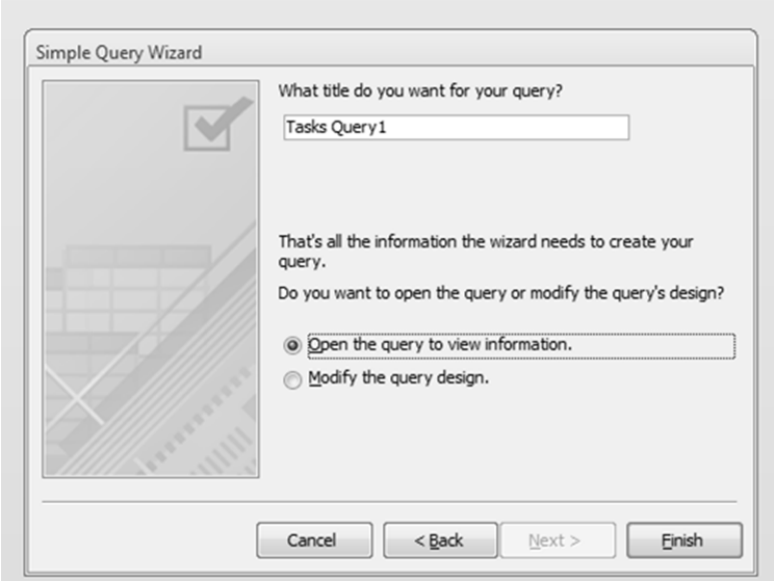


Figure 2.28: The simple query wizard

And after pressing "finish" the query is created and executed. See Figure 2.29 for an example.

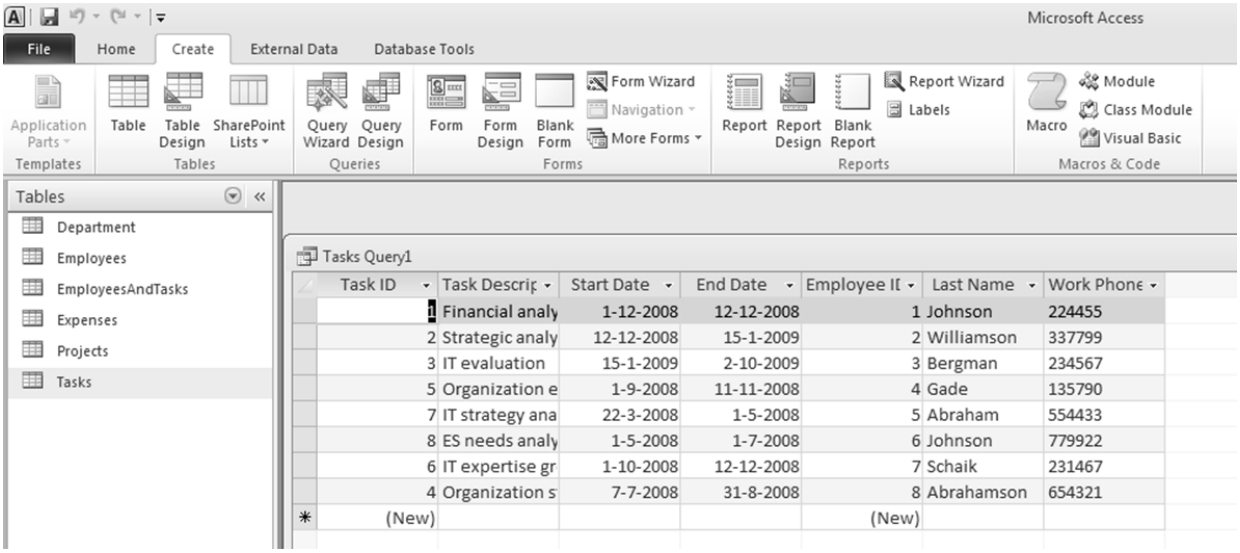


Figure 2.29 The executed query

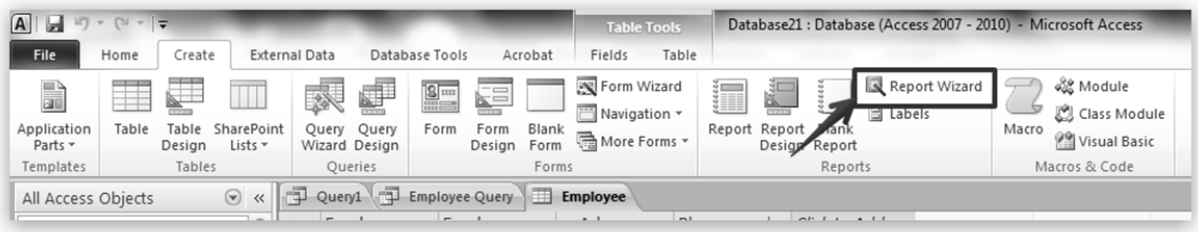


Figure 2.30 Report Wizard button

Now that you know how to create a query, we can continue with creating a report. For this, click the report wizard button. You will get a series of screens similar to the query wizard. A lot of additional report formatting options exist, which are a matter of taste and something for the reader to explore. When finished with the report wizard, you will get to the report design screen in which you can move the elements of the report around.

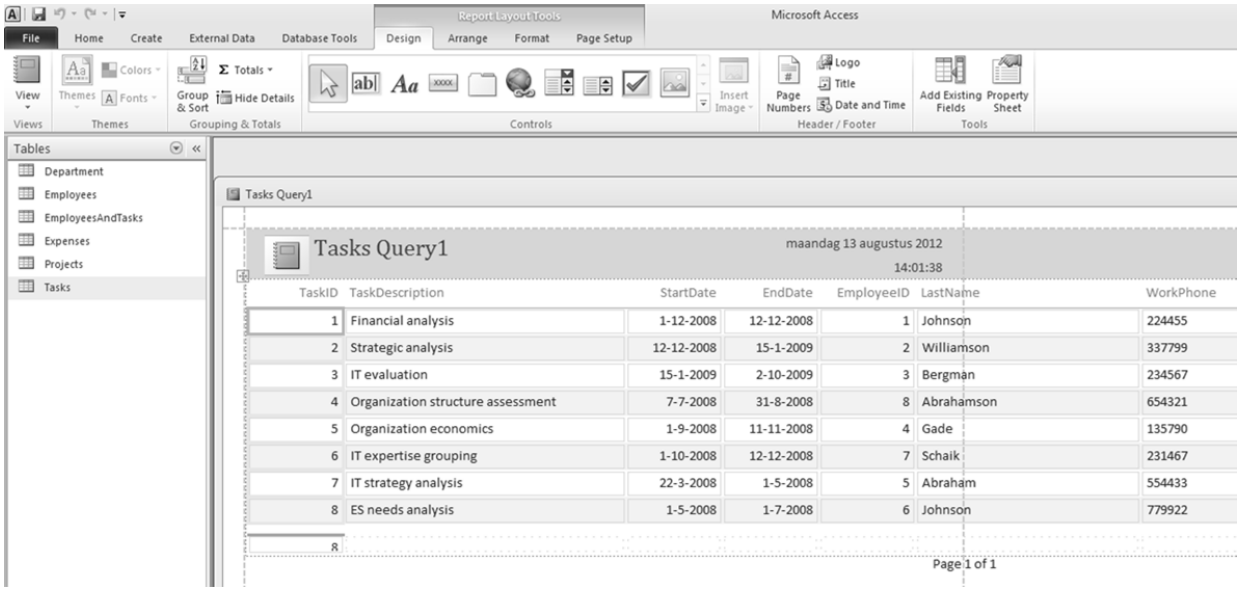


Figure 2.31: The report design screen

Now we are finished exploring MS-Access. This assignment consisted of two parts: implementing a simple database design and making a report which answers a certain problem or information need. Moreover, this is only a small part of Access, the program offers a lot of other things.

The Access file must meet the following requirements:

- For all entities from the database design, tables have been created, and for all attributes from the database design, fields are made in Access;
- The primary keys are the same as the keys in the database design, as well as the relations between entities (tables) and thus the foreign keys;
- For each attribute data is filled in, otherwise there cannot be a report;
- There is a report where attributes from various entities are linked (in the example you see for example that the employee belongs to a certain task);
- The report is clear, and you understand yourself what the report means.

2.8 Further study

This chapter introduced elementary concepts of database software. Database management is a separate profession and the database industry is extensive. For further study we recommend e.g. J. Hoffer, M. Prescott and F. McFadden (Hoffer et al., 2007). For a quick overview of alternative approaches on data modeling see (Kim & March, 1995).

2.9 Exercises

1. Create a database for a book shop.

   Things you have to do for this assignment:

   1. Set up the UoD and UoDD in MS Visio. Argue which relationships you have used and why.
      - We use the entities: Purchase staff, sales staff, products in inventory, sales reports, client card information
      - Purchase bases its purchases on client card information and sales reports
      - Purchases go in inventory first
      - Sales staff makes use of sales reports of the past and client card information to efficiently sale the books
      - For the attributes you can make up 2 attributes for each entity, with at least one identifier per entity.
      - Note that you mention the foreign keys in the UoDD.
   2. Design the database
   3. Realize the database in MS Access
   4. Create a report:

Suppose the Logistics Manager has a question with regard to the inventory and sales reports. Make a report to provide the Logistics Managers with the relevant information.

2. Create a database on a topic of your preference. If you lack inspiration, do it for a football club (teams, players, coaches, team leaders, support, administration, maintenance) or for a flower shop (suppliers, sales staff, inventory, sales reports, and client card information).
    1. Set up the UoD and UoDD using Visio.
    2. Design the database
    3. Realize the database in MS Access
    4. Create reports

    Do this job preferably with at least one other person.

3. Review question: What is the difference between data managed via a structured database and data managed via social software?
4. Key question: Why, when and how can databases enable a collective brain? What is the content of this collective brain? What processes and rules enable that this brain may be called collective?

3. Rationalism and spreadsheets for decision support

3.1 Leibniz and spreadsheets

According to 17th and 18th century rationalist philosopher Leibniz there are two kinds of truths (Huenemann, 2008). The "truth of reasoning" and the "truth of fact". "Truth of fact" is the empiricist perspective discussed in the previous chapter on databases. The empiricist or Lockean inquiring system approaches information management as the management of representations of observations and experiences non basis of which one can know the status of something or at best can generate correlations among the values of attributes of entities. According to Leibniz, this "truth of fact" is far below what people are capable of. Following our understanding of Leibniz, what people really need is explanations and predictions so that they know what happens if certain actions would be taken (Huenemann, 2008). Leibniz states that if we are not able to produce such explanatory or predictive insights, we will not be much more effective than dogs, who do understand regularities like performing certain tricks for earning sausages. People are not dogs, and the main thing that distinguishes people from animals is our reasoning capability, that is the capability of understanding if-then or what-if relations, which are causal relations gained by reflections about past events (i.e, explanations) or reflections about what could happen in the future when something changes (i.e., predictions). Causal knowledge is obviously a specific kind of knowing that can be well represented by causal models and thus becomes information and enables the production of new decision relevant information.

As one may probably conclude from the statements above, the difference between rationalism and empiricism is not that black-white Huenemann (Huenemann, 2008). Empiricists do use causality and reasoning in their view on knowledge creation and rationalists prefer empirically tested models over those that have no connection to any reality. So for example, empiricist Hume (Norton, 1999) states several empirical preconditions for the perception of causality:

1. The cause and effect must be contiguous in space and time, i.e., the "if" goes before the "then".
2. The cause must be prior to the effect.
3. There must be a constant union between the cause and effect.
4. The same cause always produces the same effect, and the same effect never arises but from the same cause, or at least we perceive it like that.
5. Where several different objects produce the same effect, it must be by means of some quality, which we discover to be common among them.

6. The difference in the effects of two resembling objects must proceed from that particular, in which they differ.
7. When any object increases or diminishes with the increase or diminution of its cause, this is to be regarded as a compounded effect, derived from the union of the several different effects, which arise from the several different parts of the cause.
8. An object, which exists for any time in its full perfection without any effect, is not the sole cause of that effect, but requires to be assisted by some other principle, which may forward its influence and operation.

The key point for rationalists however is that some important knowledge is not necessarily empirical because it gives elementary abilities to reason, like in mathematics and logic. In this section, when we use the term "causality" we mean this in a rationalist sense, i.e., as any kind of "if then" or "what if" statement, empirical or logical.

This view on truth leads to a number of implications for information management. For example: computers cannot make logical connections without people who actually "know" (or *think*) causal relationships. That means it is essential to point out causal relationships before computers can help us with reasoning and decision making. Data in the empirical sense may be input and output of causal models. Poor data input, may result in garbage output (the so-called garbage-in, garbage-out or GIGA effect), but sometimes reliable input data do not exist and still reasoning under given assumptions (named simulation) may be very useful for well-reasoned decision making (especially in strategic decision making which has to reason about not (yet) existing futures which are often names scenarios (Schoemaker, 1995)). Causal models have a reasoning mechanism to infer impacts of states of substances on the states of other substances. Spreadsheets are very suitable for implementing causal models for analyzing the impacts of certain inputs on certain outputs as we will explain in this section. The inputs of models and spreadsheets can be representations of future objects (like profits) and the model may reason back to the required means (which are the model's output) to achieve them, or the inputs may be representations of existing conditions (like the number of clients or the existing financial resources) by which certain decisions may result in certain impacts in the longer run (and thus are the model's outputs).

For going from problem scopes to effective use of software (MS Excel in this case), we apply a variant of the general model of informing introduced in chapter one, named the rationalist model for decision support (see Figure 3.1).

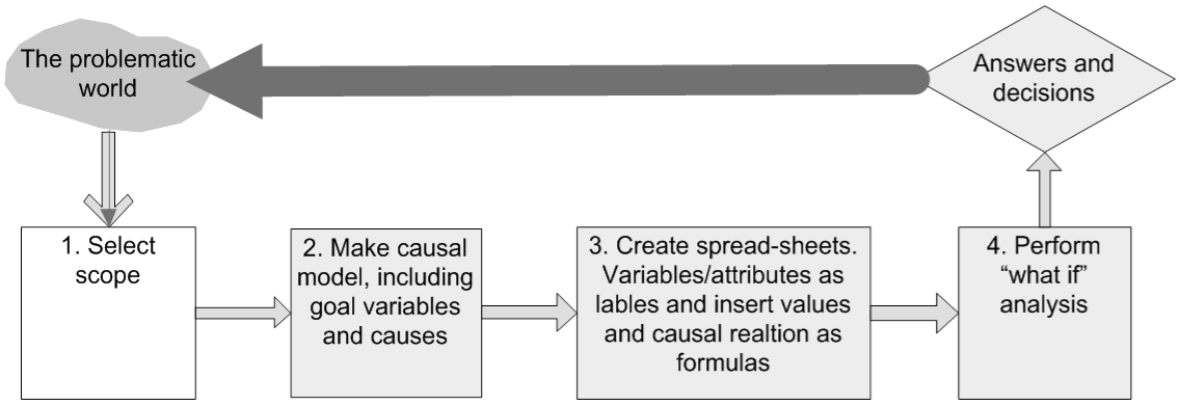


Figure 3.1: The rationalist model for decision support

3.2 Goal of this chapter

Following our understanding of Leibniz, what people need are explanations and predictive models so that we know what happens if we take certain actions. This we can achieve through the making of causal relationships and implementing them in the automatic reasoning tool MS Excel. MS Excel also has many excellent tools for analyzing data that are a good addition to a database. Therefore this chapter will present the following:

1. First a description of MS-Excel and its opportunities to perform descriptive statistics and data visualization in section 3.3. Note that this MS-Excel description is based on the UK 2013 MS-Excel version. Some notes on different languages are mentioned in chapter 1. The Microsoft corporation is reasonably consistent in its interfaces and user options in versions from 2003 onwards, although some items may be a bit different of you have a non 2013 version.
2. Second we explain how one can represent causal reasoning into a causal model by using MS-Visio in section 3.4.
3. Third, we explain how one can implement causals models in MS-Excel in section 3.5 so that in MS-Excel one can work with and make calculations using the causal structures identified.
4. Finally, we will demonstrate how MS-Excel can handle uncertain causalities and thus decisional uncertainties and the lack of data in causal reasoning.

Ultimate goal: Understanding of the Leibnizian view of information as causal models that can be processed by computers to speed up reasoning and decision making.

3.3 Elements of spreadsheets

A spreadsheet is a table with columns (labeled by letters in alphabetic order) and rows (labeled by numbers), which allows to perform calculations and reasoning. The input of data in a spreadsheet is

realized by putting data in the cells. These are just values (not attributes) that do not have a function other than input data. Causal relationships give the data a function and make sure that the cell has a goal as a representation of the value of an independent variable or dependent or goals variable in a causal model. For example, the causal relation makes a connection between the revenue variable on the one hand and the expense variable on the other hand. This is what makes a profit or loss value. The related value is calculated on other primary data by a formula and thus the output data of the model is called "derivative" data.

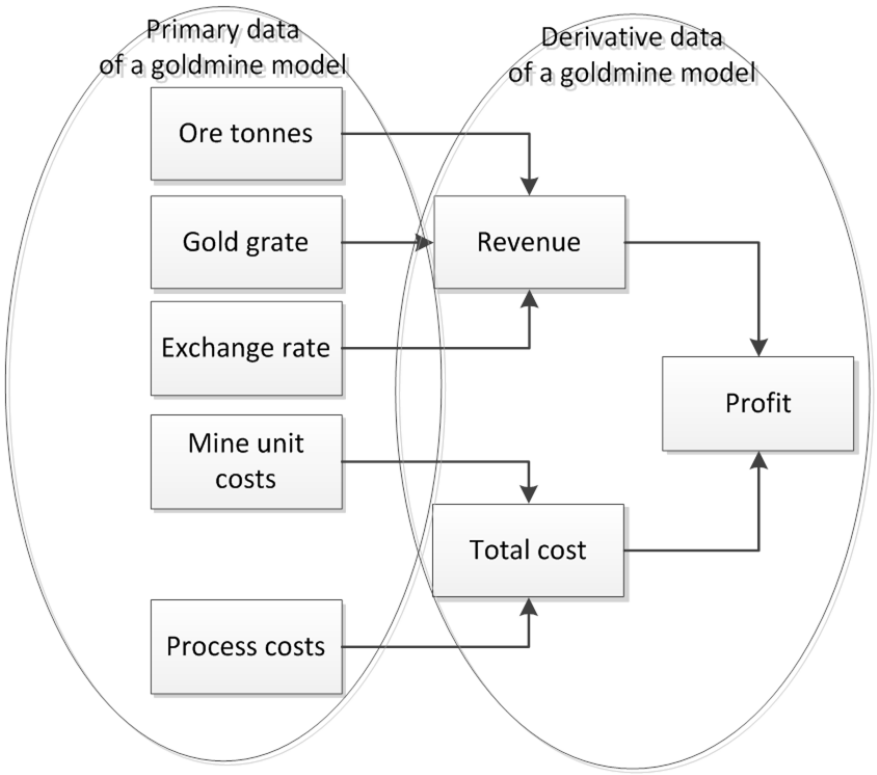


Figure 3.2: Primary and derivative data in a causal model

MS-Excel works with books which wear the extension '.xls' or "xlsx" just like a Word document wears the extension '.doc'. An MS-Excel book consists of one or more tabs. MS-Excel names these tabs automatically Sheet1, Sheet2 and Sheet3, but these can be renamed by clicking on the right button of your mouse and choose 'Rename'. You can also add or remove Sheets by clicking 'Insert' or 'Delete' in the menu named earlier. A Sheet exists of rows (1, 2, 3,…) and columns (A,B,…,AA, AB,…). See Figure 3.3.

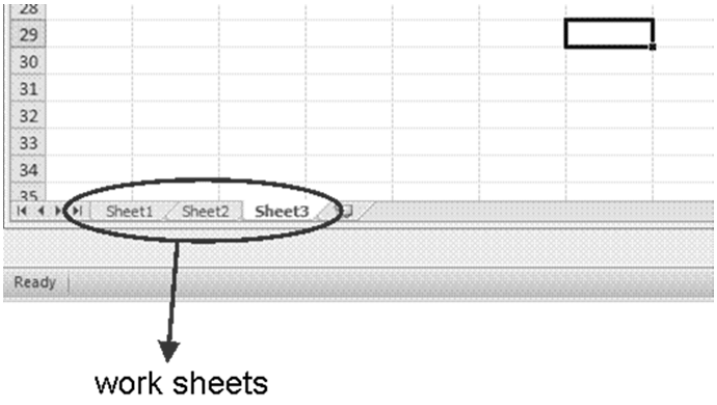
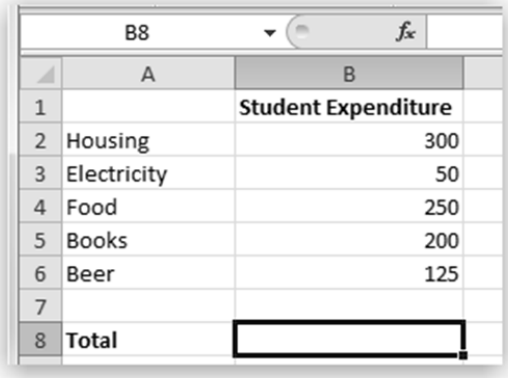


Fig. 3.7: The MS-Excel work book

One of the basic advantages of a spreadsheet over the (graphics) calculator is that you can set up a complete computation scheme in which you convert input data to output (results). Once you have such a scheme, you can modify input values to obtain modified results instantaneously.

If you need to dynamically calculate values that are dependent on other values entered in your worksheets, then you need formulas. To create a formula that uses data in certain cells, you need to refer to those cells. MS-Excel uses a special notation to point to a particular cell; first the column letter and then the row number are used. For example to refer to the cell in the second column (column B) and the third row (3) this would be cell B3, which is the expenditure on electricity in the case of figure 3.4. In these cells all kind of standard arithmetic expressions can used, like for Addition: +; Subtraction: -; Multiplication: *; Division: /; Power: ^, for example: "=2^3" and gives 8 (2^3=8).



	A	B
1		Student Expenditure
2	Housing	300
3	Electricity	50
4	Food	250
5	Books	200
6	Beer	125
7		
8	Total	

Fig. 3.4: Variable labels (column A and B1) and data (B2:B6) in MS Excel

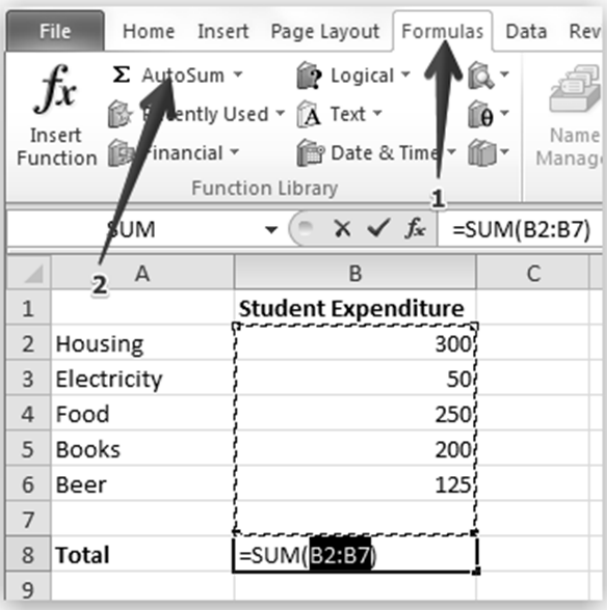


Fig. 3.5: The formula bar (B) in MS Excel

To get started we demonstrate the workings of the "Sum" formula in figure 3.4 and 3.5. As you can see we have made an overview of several of the items on which a student has to spend money each month. What we want to do is to get the total of this list of items. To do this we use the AutoSum feature of MS-Excel. Place the cursor at cell B8, which is the empty cell in the column next to "total". Then click on the "Formulas" tab as shown in Figure 3.5. Next click on the AutoSum button. MS-Excel automatically guesses the range of cells that you want to sum and puts it in the Sum formula for you. Press Enter and MS-Excel will calculate the total, and thus produces derivate data.

Note: It is important that you understand how to build a formula from this exercise:

- First: you can see that the formula starts with "=". In order for MS-Excel to recognize something as a function it MUST start with "=".
- After the "=" you will state the name of your function, which is SUM in this case, since we want to sum data.
- Next you start with an open bracket "("
- Between the brackets you will give the ARGUMENTS of the function. Each function has different arguments which you can find in the Help function of MS-Excel (press F1 for Help).

- End with a closing bracket “)”

Since we want to sum data, we use the SUM function; our argument is the range of the cells that we want to sum, which is Cell B2 to B6 as you can see in figure 3.4. The AutoSum feature also includes the empty cell B7, this is no problem since MS-Excel just ignores the empty cell. The final formula therefore is “=SUM(B2:B7)”.

We continue with another example. We have the price of different items and we want to calculate the total price of the item including VAT (BTW in Dutch). We first calculate the VAT on each item and then add it to the price to get the total price of each item. Let’s calculate the VAT first, based on a VAT rate of 21%. The formula therefore is VAT = 21% x Price. Make cell B2 active and in it type =21%*A2 (see Figure 3.6). MS-Excel knows how to handle percentages so you don’t need to convert 21% to 0.21 for the calculation. When you press enter, MS-Excel calculates the VAT to be €31.50. Now that we have calculated the VAT for the first item, we can use this formula for all the other cells. Instead of typing the same formula for each item, there is a quicker solution. Select Cell B2 which contains your formula, what you see is that in the lower right corner there is a “dot”. Hoover your cursor over the dot and you will see that the mouse cursor changes to a small cross. Next, drag the cursor downwards to select all cells that you want to paste the formula in (which is down to cell B6 in this case). Release the cursor and MS-Excel will copy the formula to all the highlighted cells.

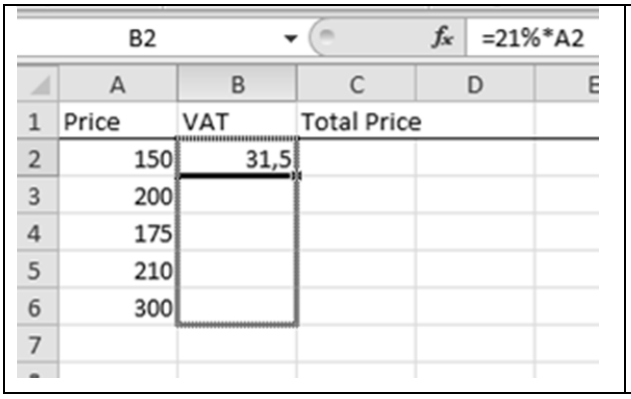


Fig. 3.6: Formula drag down

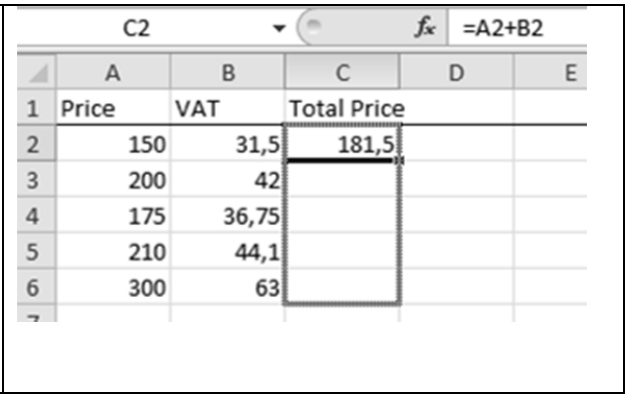


Fig. 3.7: Formula drag down

If more prices have to be added, one may also insert a row by right-clicking on the left number of the spreadsheet and apply “insert row” above or below. In a similar same way columns can be added (for example “product names” before the “Price” column.

Obviously, Total price = price + VAT, which we can implement in the spreadsheet by typing formula =A2+B2 into cell C2. When you press Enter, MS-Excel calculates the sum of A2 and B2 to give 181,50. We next can again apply "drag down" to copy the formula of C2 down to C6 and get the total price for all the items. Note that intentionally I keep the comma decimal separator, which is common usage in for example German, French and Dutch versions of MS-Excel, where in English versions point separators will be used. This means that it is important to know what language version of MS-Excel you have. A point or comma mostly mean something totally different.

One may also put the VAT % in a separate cell, for example in cell C2 and from there calculate B2:B6 via the formula "=A2*C2" in B2 and dragging this formula to B6. However, this results into errors for cells B3:B6 because B3 will change the formula to A3*C3, whereas C3 is empty. This happens because MS Excel uses relative cell addressing by default. To solve the problem you need absolute cell addressing by putting a $ sign in front of C2. Therefore, the formula of B3 changes to A3*$C$2 and thus the percentage mentioned in C2 can be reused. Absolute cell addressing is an important option for enabling reasoning with data, as we will explain later in section 3.5.

It is also possible to refer to cells of different worksheets in another worksheet. For example: add cell A1 of Sheet 2 (the profit of last year) to Sheet1 containing the calculations for this year. You can do this by referring to the cell as [Sheet Name]![Cell]. So in this case the formula would be =Sheet2!A. The answer now appears in the cell. See Figure 3.8.

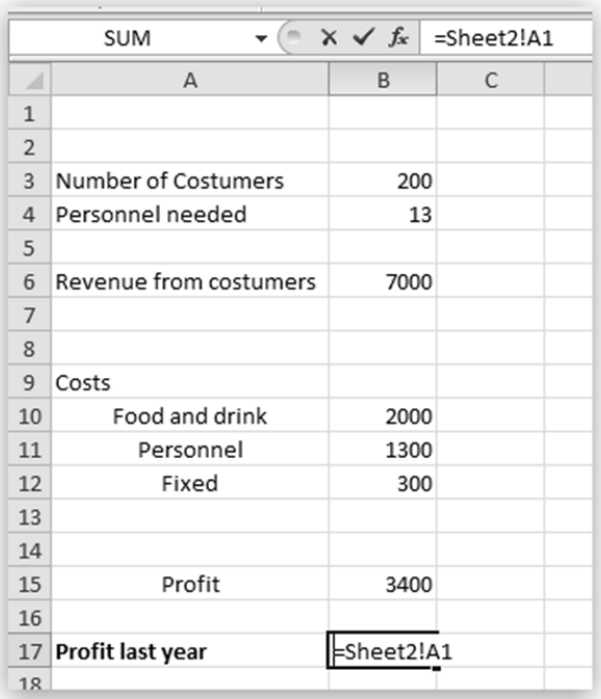


Fig. 3.8: Taken data from other sheets in a formula

The SUM formula we have used so far is an easy formula, but sometimes you might not know how to use a certain function or perhaps don't even know the name of it! This is where the Insert Function option of MS-Excel comes in handy. This feature helps you to find the function you need and provides feedback as to which arguments the function needs.

When you click the button as shown in Figure 3.9, the Insert Function screen will show. In this window you can type a description of what the function needs to do and MS-Excel will help you and find the right function for your needs. Of course you can also browse the list and click on functions to see their workings. See for example Figure 3.10, here you can see that the Sum function is selected. MS-Excel displays the name of the function, the arguments that it takes and an explanation of the workings of the functions ("Adds all the numbers in a range of cells"). If this is not enough information for you, you can click on the "Help on this function" link and the MS-Excel help will open which is more detailed and should get you on the right track.

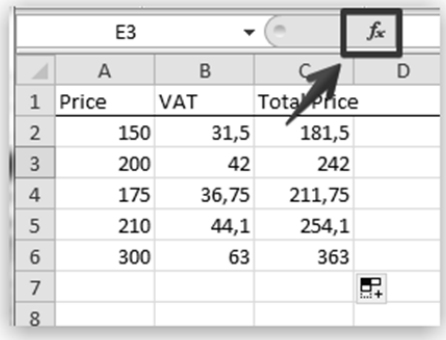


Fig. 3.9: Insert function

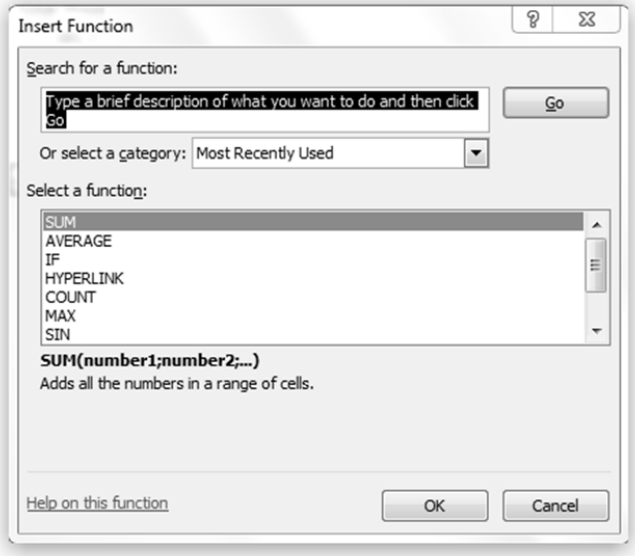


Fig. 3.10: Select a function from the list

MS-Excel comes with many built in functions that cover a wide range of topics. Some of the more commonly used ones are given below.

- ABS: Returns the absolute value of a number
- AVERAGE: Adds its arguments
- COUNT Counts how many numbers are in the list of arguments
- MAX: Returns the maximum value in a list of arguments
- MIN: Returns the minimum value in a list of arguments
- SQRT: takes the square root of its argument
- ROUND: Rounds the argument to the nearest integer.

See http://www.excel-2010.com/excel-function/ for more information.

Giving arguments to a function

Some more examples of different ways to reference to cells as an argument to a function are given in Table 3.1. We have taken the SUM function over a range of values. The data of Figure 3.11 has been used as input for the formula.

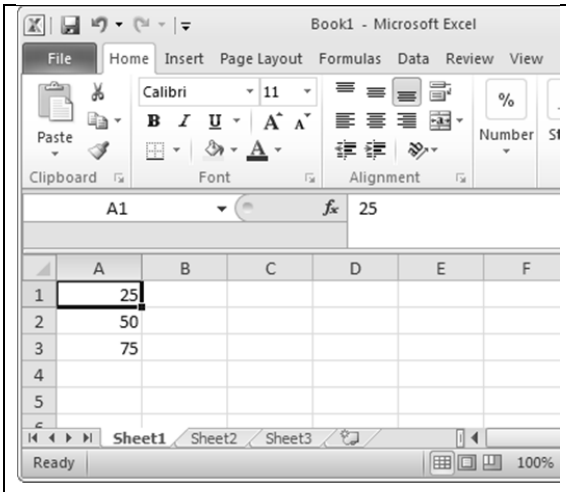


Fig. 3.11: A simple sheet

Example	Cells to add	Answer
=SUM(A1:A3)	A1, A2, A3	150
= SUM(A1:A3; 100)	A1, A2, A3 and 100	250
= SUM(A2+A3) or: = SUM(A2;A3)	A2, A3	125
= SUM(A1:A2; A5)	A1, A2, A5	75

Table 3.1: Explanation of actions on Sum functions for Figure 3.11.

Advanced functions - The Net Present Value formula

A very common and useful MS Excel function is net present value calculation, for which we give the example of a 200,000 euros investment, which discounts over 5 years and for which revenue forecasts are given. The NPV function has the following structure: First the discount or interest rate, next the revenues that are expected. From this we subtract the initial investment costs (Cell C4).

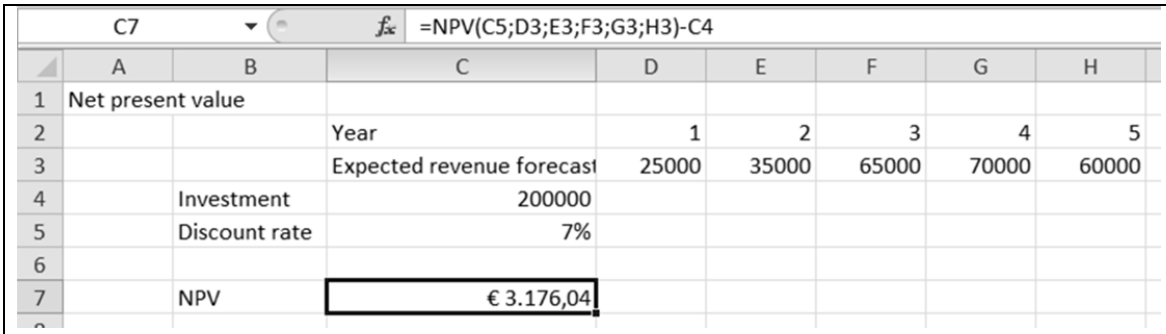


Figure 3.12: Net present value calculation

There are many other advanced functions that make life easy for you. Some examples;

- STDEV(): gives the standard deviation of the given cells in the range.
- VAR(): gives the variance of the given cells in the range.

- PRODUCT(): returns the product of the given cells in the range. Example: PRODUCT(A1:A3) returns the value of A1*A2*A3
- SUMSQ():Returns the sum of the squares of the arguments. Example: SUMSQ(A1:A3) returns the value of $A1^2+A2^2+A3^2$
- SUMPRODUCT(): Multiplies corresponding components in the given arrays, and returns the sum of those products. The arrays should have the same length. Example: SUMPRODUCT(A1:A3; B1:B3) returns the value of A1*B1+A2*B2+A3*B3.
- CORREL(array1; array2): Returns the correlation coefficient of the array1 and array2 cell ranges. Use the correlation coefficient to determine the relationship between two properties. For example, you can examine the relationship between a location's average temperature and the use of air conditioners.

Common Error announcements in MS-Excel

In some cases, entering a formula does not yield a result that you expected. You may see that a cell contains the indications like #VALUE, #NUM or #DIV. What does this mean? In general, your formula is wrong or it refers to cells that contain unexpected values. To be more specific:

- #DIV/0: Occurs when a number is divided either by zero (0) or by a cell that contains no value
- #VALUE: Occurs when the wrong type of argument is used (for example, a character instead of a number)
- #REF: Occurs when a cell reference is not valid, for example if you refer to a cell in a nonexisting work sheet.
- #NUM: Occurs with invalid numeric values in a formula or function, for example if your try to compute the square root of a negative number like "=SQRT(-1)"
- #N/A: Occurs when a value is not available to a function or formula
- ########: Excel generally displays this error when a column is not wide enough to display all the characters in a cell. You can solve this problem by adjusting the column width (see below), or by changing the numerical format (see next section)

What can you do about it? It may be helpful to click the cell that displays the error, click the button that appears  , and then click Trace Error if it appears. With respect to the last error (########), here is a simple way to change the column width. Move the arrow to the right side of the column label and click and drag the mouse to the right (to make wider) or left (to make smaller). Let go of the mouse button when the column is wide enough. 

Notice the cursor changes to a vertical

line with arrows pointing left and right. You can double-click that vertical line (see image) to auto-size the column to the cell with the maximum width. You can also change the vertical height of a row by moving the lower edge of the row title (number).

Logical Functions

Also, Excel contains handy LOGICAL FUNCTIONS. For example, the IF-function checks the logical condition of a statement and return one value if true and a different value if false. The syntax is: "=IF (condition, value-if-true, value-if-false)", which is a logical function that returns one value of the condition is true and another value if the condition is false. For example, when students pass an exam (P) when their score is 6 or higher, and fail (F) when the score is less than 6, a teacher could apply in his spreadsheet the function "=IF(A2<6, "F", "P")".

IF-functions can also be nested by replacing a "value if true" or "value if false" statement by another IF function, like "=IF(SUM(A1:A10)=5, IF(B3<3,1,0), "")" for cell C3, which actually says: If the sum of cells A1 till A10 is exactly 5, check if the value of B3 is less than 3 and produce a 1 in cell C3 or else produce a 0 in C3 if B3 is equivalent to 3 or larger. If the sum of A1:A10 is not equal to 5, produce nothing in C3. Note that equal and larger or smaller is presented by =< and => and that "not" equivalent to is presented by < >.

Conditions can be combined with the AND() and OR() functions.

MS-Excel Data Formats

When you enter data into a cell in MS-Excel, it is handled differently depending on what type of formatting you have assigned to the cell. Here is a screen displaying what you would see if you select a cell (or group of cells) and from the right-mouse-click-menu "Format Cells" (alternatively, press CTRL-1):

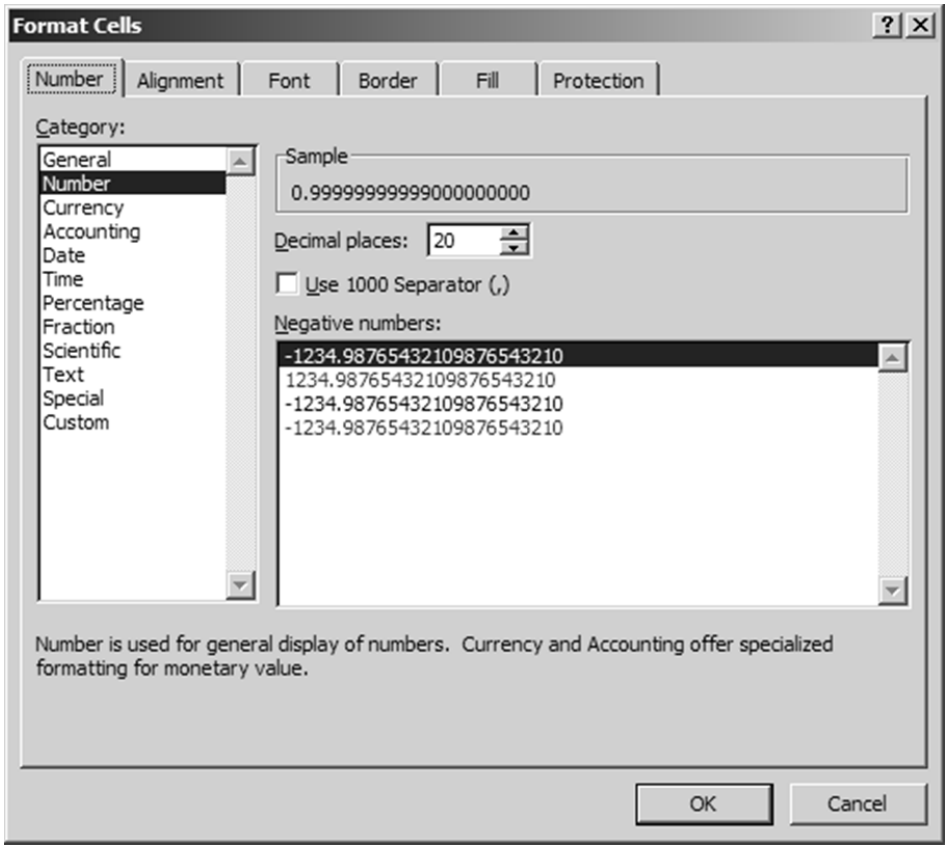


Fig. 3.13: MS-Excel data formats

For example, if you choose Currency formatting, MS-Excel automatically converts 3 into €3.00; if you choose Date formatting, MS-Excel converts 3/1 to March 1, 2002; and if you choose Percent formatting, MS-Excel changes 0.3 to 30%. There is also the Special option in MS-Excel, by which you can select formats like Social Security Number and Phone Number.

If you have a large spreadsheet containing headers with information on the cell contents, it is useful if you can always see those headers, even if you scroll down the data list. You can easily do that using Freeze panes as shown in Figure 3.14.

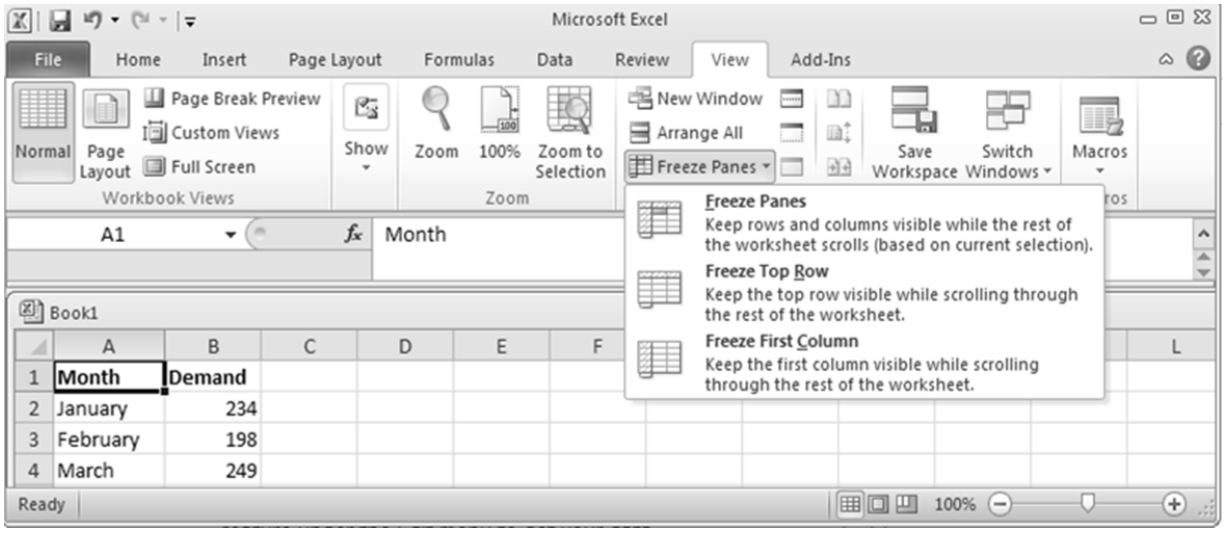


Figure 3.14: MS-Excel's Freeze function

In MS-Excel you can give numbers a continual layout. This happens when you select the cell and press the right button of your mouse, and choose the option 'Format Cells'. With this screen you can edit the features of the cell, like the font, color, thickness of the borders. See Figure 3.15.

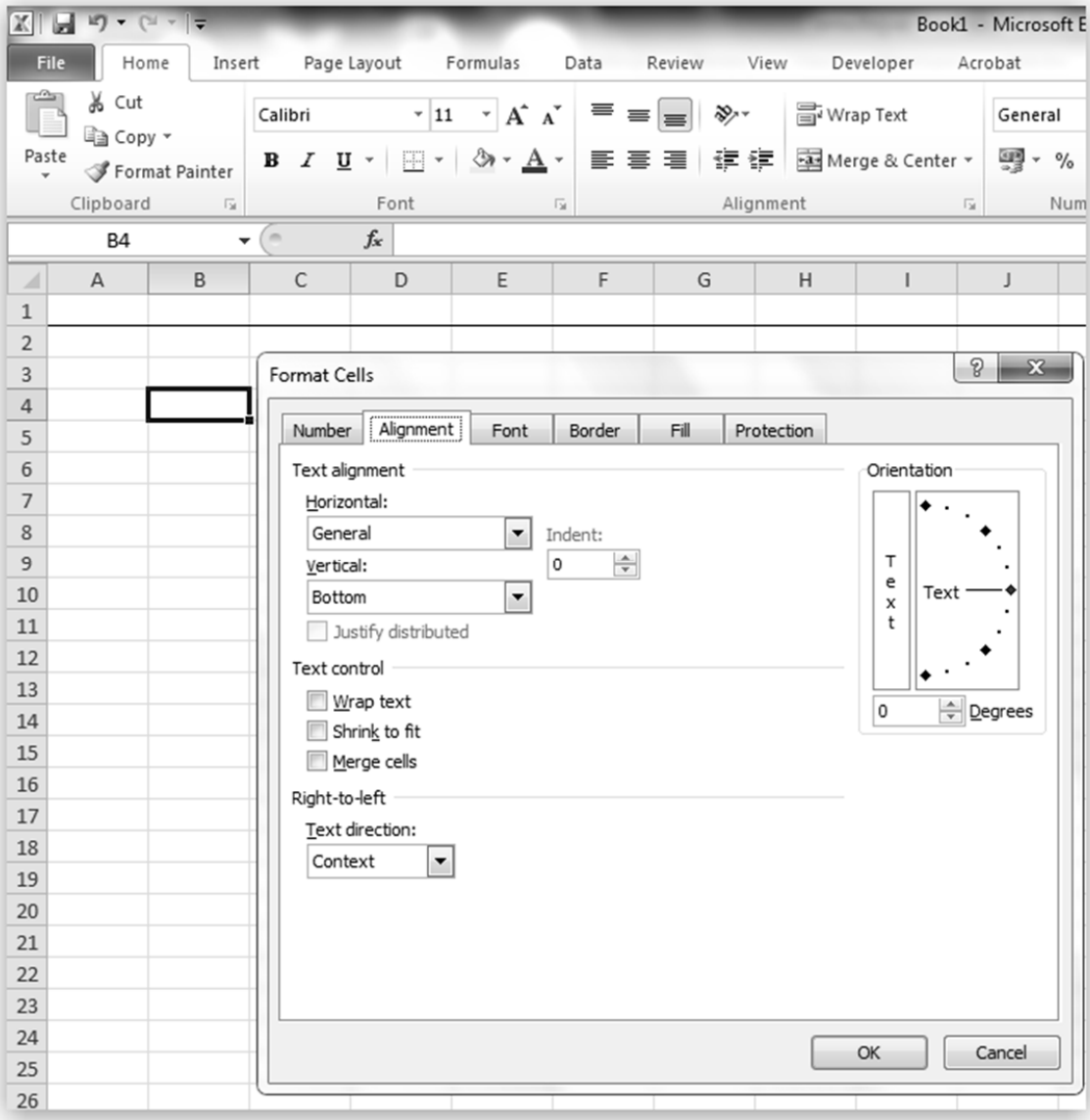


Figure 3.15: Cell format selection.

You can also automatically adjust the cell size to the amount of text in the cell by under ‘Format cells’, the tool: ‘Alignment’.

3.4 Producing descriptive statistics and data views.

Databases are useful for persistent and structured storage of large volumes of data. MS-Excel is useful to analyze the data, create charts, perform various calculations, what-if analyses etc. There are various ways

to work with database data in MS-Excel. For achieving this check, your MS-Excel help file Importing data from Access for an overview. The preferred way is to create a connection to the MS-Access query containing all merged data. Updates in your database will then be reflected in your MS-Excel datasheet. However, note that you cannot create a connection when the database is still open in MS-Access. You can create a one time export in MS-Access of the query data and import the data in MS-Excel but note that all updates to the MS-Acces database will not be incorporated in the MS-Excel file.

Create the connection in MS-Excel (Data Connection) as shown in Figure 3.16.

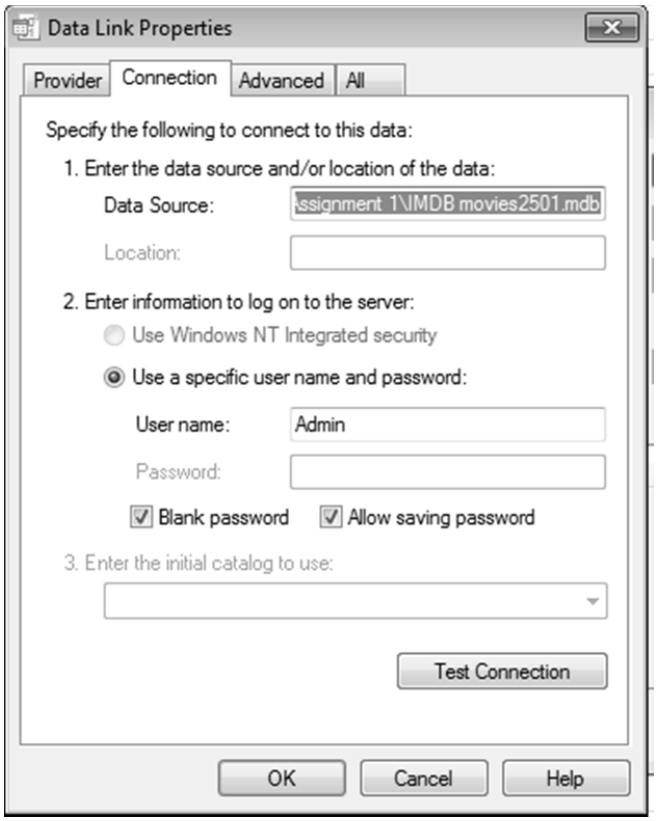


Figure 3.16: Data link with access

If the connection is successful, you will be allowed to insert data through an existing connection. If your version of MS-Excel does not support this simple import, you also can export the Query data from MS-Access to an MS-Excel file.

Means and standard deviations can be easily calculated in MS-Excel by selecting the proper list of data and the formula. With frequencies this is a bit more complex, because it requires new tables being

produced from calculations of another table. Frequency tables and correlations both are not just formulas but combinations of different data.

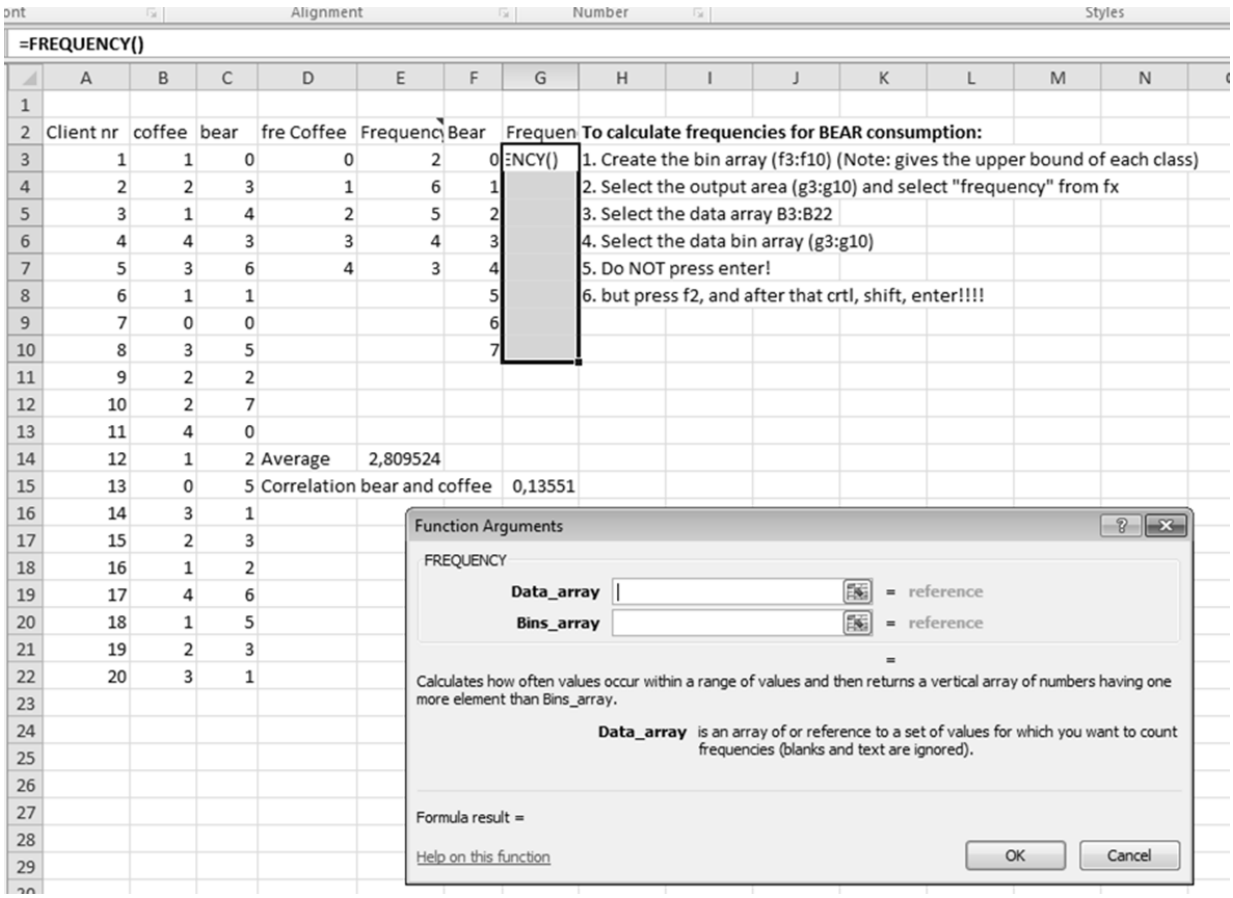


Figure 3.17. Calculating frequencies for Beer consumption.

Use the "=round" function if you want to the upper bound with rounded numbers.

Typically data appears in flat tables, i.e., data that only consists of rows and columns. When the number of rows gets very large it can become difficult to get summarized information from the data. This is where pivot tables come in handy. A pivot table can help to summarize data and show the data that is most important to the user. Pivot tables can be used in many situations and the way they are used varies with every situation. For an example on pivot tables and how to convert tables to pivot tables see Figure 3.18.

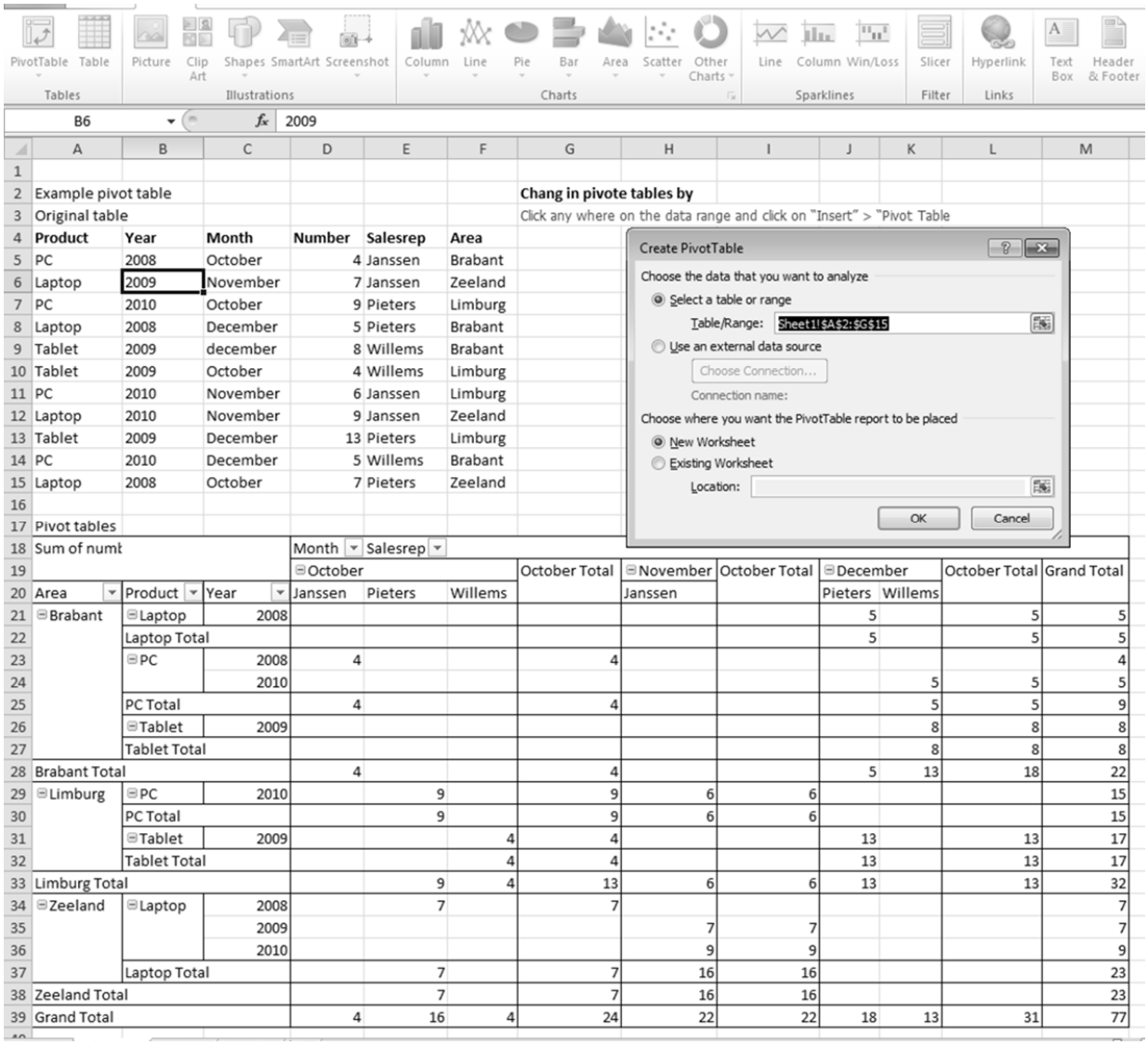


Figure 3.18. Pivot table example

A screen for selecting column and row labels and values helps in quickly customizing the pivot tables. See Figure 3.19 and 3.20.

19	Sum of Number	Column Labels			
20	Row Labels	Janssen	Pieters	Willems	Grand Total
21	**Laptop**	**16**	**12**		**28**
22	November	16			16
23	December		5		5
24	October		7		7
25	**PC**	**10**	**9**	**5**	**24**
26	November	6			6
27	December			5	5
28	October	4	9		13
29	**Tablet**		**13**	**12**	**25**
30	December		13	8	21
31	October			4	4
32	**Grand Total**	**26**	**34**	**17**	**77**

Fig 3.19: The table labels

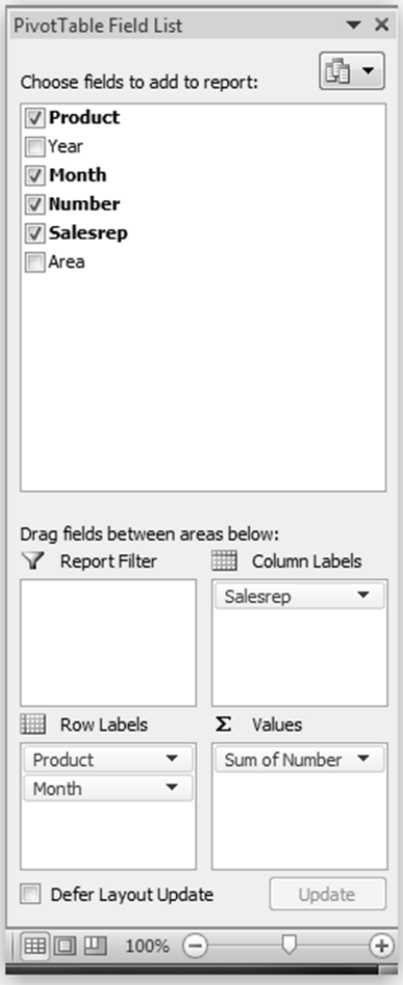


Fig. 3.20: And the related Pivot table field list

When you select fields to show in the pivot table MS-Excel looks at whether the data is numeric, in this case it will appear in the values section (see the "Sum of Number" sold items in example Figure 3.20). When the data is found to be a string, like the product or month in our case, it will fall under the heading of "Row Labels". By using different ways to show the pivot table you can get a better understanding of large portions of data and it will support you in making the right decisions based on this data.

The menu function "Data Sort" enables you to quickly sort a selected table with data. The function allows you to use multiple sorting criteria (e.g. first sort on column D, then sort all ties on column B, and finally

sort all remaining ties on column A). If you have defined headers for each column, you can indicate “My data range has headers”. Now you can select the headers (instead of column A, B, etc.) as sorting criteria. You can sort ascending or descending.

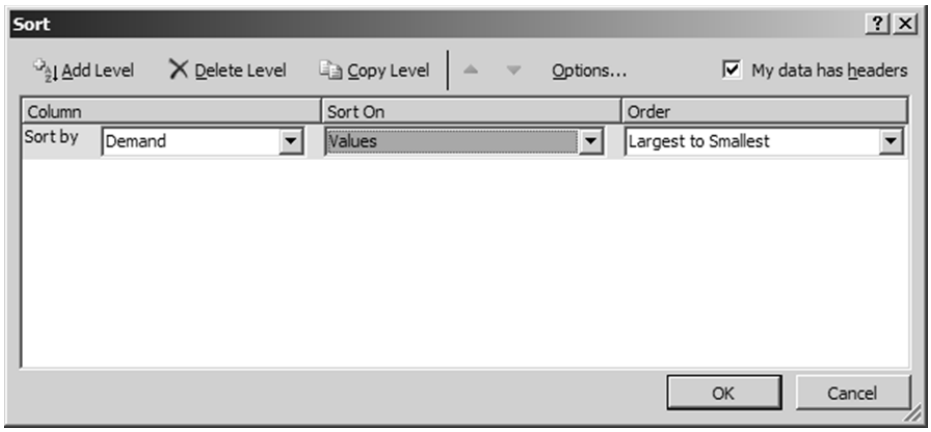


Fig. 3.21: Data sort

Graphs and charts are a good way to visually represent data. MS-Excel offers most common chart types and makes it easy for the user to compile charts quickly. To create a chart, select the data that you want to include in the chart, click on the insert tab and select the kind of chart that you would like.

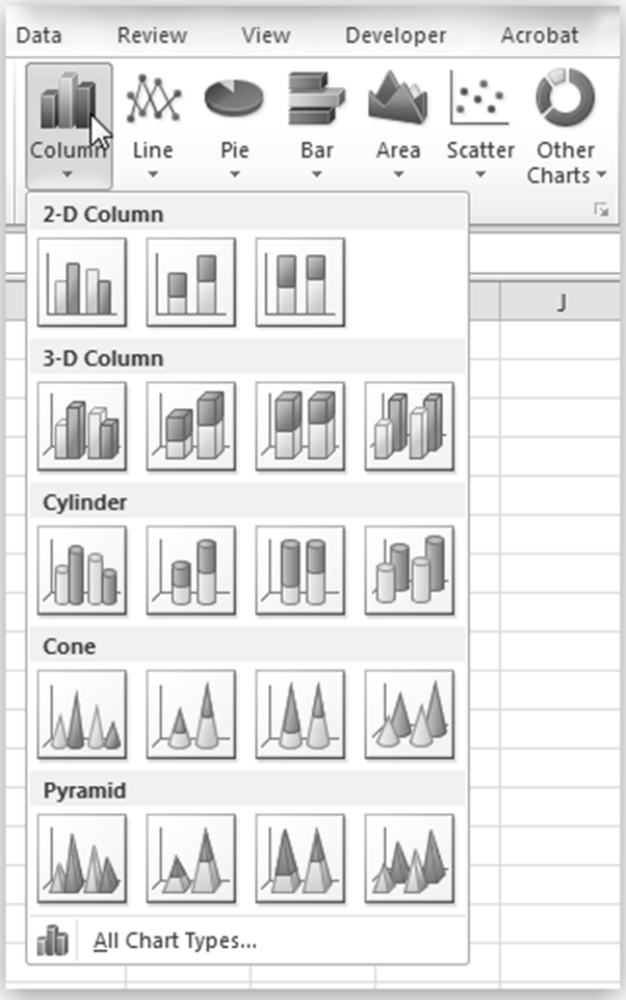


Fig. 3.22: Chart types in Excel

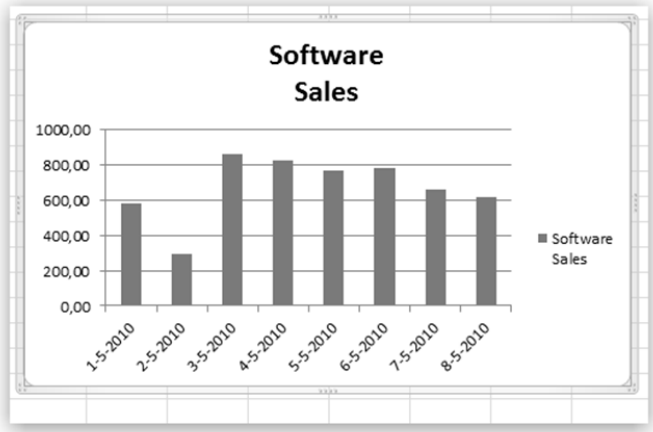


Fig. 3.23: Chart attributes

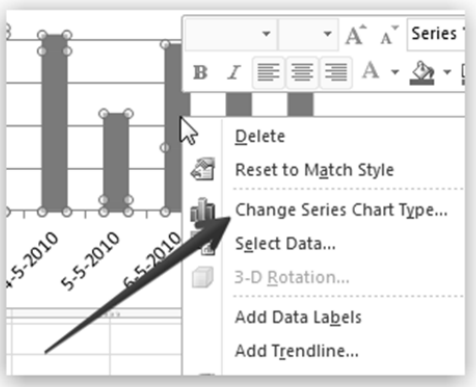


Fig. 3.24: Changing chart attributes

The Column chart is one of the most commonly used chart type and is used to show the changes in data over a period of time or illustrate comparisons among items, see Figure 3.23. When the chart is shown, it is very easy to change the type of chart and attributes of the chart. Just right-click on the chart and click "change series chart type" to change the type of chart. The type of chart which is best suited varies with the type of data that you want to show. Line graphs are suitable to show changes in data over time. For example a graph of the monthly profits for the last six months (see Figure 3.25).

Fig. 3.25: Line graph

A pie chart is very useful when you want to emphasize the importance of a specific element in the data. For example in Figure 3.26 you can see that salesrep Willems has sold more Tablets than PC's in the last period.

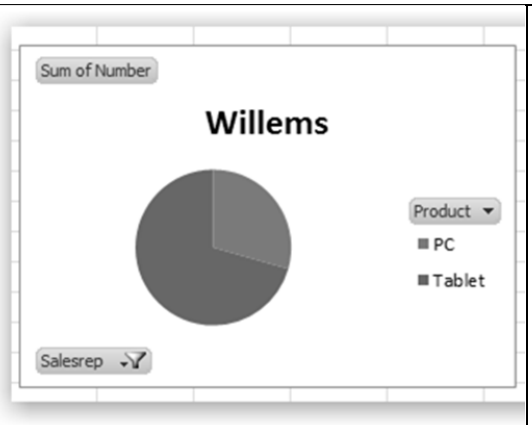


Fig. 3.26: Pie chart

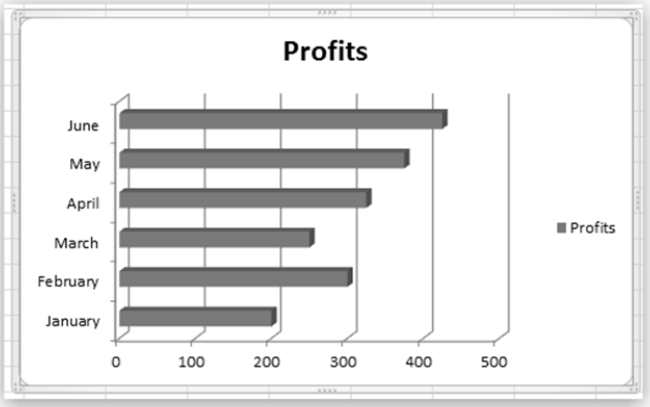


Fig. 3.27: Bar Graph

A bar graph is suited for making comparisons between items.

MS-Excel has default settings when creating a chart regarding the color of the borders, the background color etcetera. Mostly these settings are adequate, but modifying these settings is very easy. Just right click on the chart area and choose "format plot area". In this menu you can change all kind of visual settings of your chart (Fig. 3.28).

Fig. 3.28: Format the plot area

MS-Excel contains many useful probability distribution functions, such as:

- NORM.DIST(x, mu, sigma, TRUE): returns the probability that a Normal distributed random value with mean mu and standard deviation sigma is ≤ x.
- NORM.DIST(x, mu, sigma, FALSE): returns the probability density of a Normal distributed random value with mean mu and standard deviation in the point x.
- NORM.S.DIST(x, TRUE): returns the probability that a standard Normal distributed random value (with mean mu=0 and standard deviation sigma=1) is ≤ x.
- NORM.S.DIST(x, FALSE): returns the value of the standard normal density function in the point x.
- NORM.INV(p, mu, sigma): inverse of the Normal distribution function; finds the point x for a Normal distributed random value with mean mu and standard deviation sigma, such that the probability that this variable is ≤ x equals p. We give an example in the figure below. Suppose

that we have a normal distribution with mean 100 and standard deviation 20. The distribution function in the point 55.5 is given by =NORM.DIST(55.5;100;20;TRUE), which yields the value 0.013. Reversely, the inverse function gives the point at which the distribution function equals 0.013, so =NORM.INV(0.013;100;20) yields the value 55.5

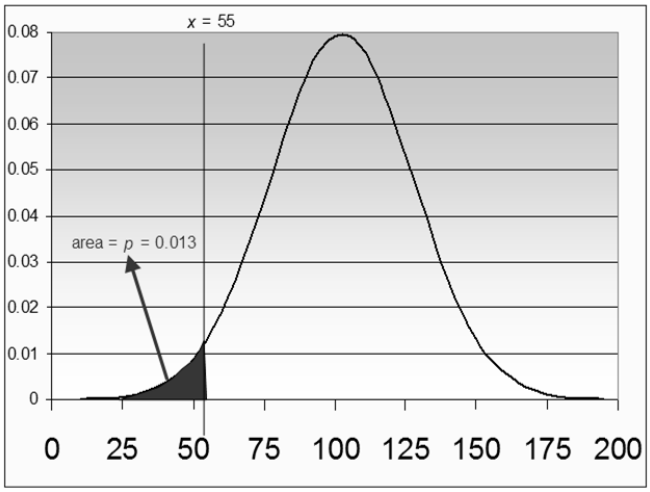


Fig. 3.29: Normal distribution

- NORM.S.INV(p): inverse of the standard Normal distribution function; returns the value x for a standard Normal distributed random value, such that the probability that this variable is ≤ x equals p.
- BINOM.DIST(x, n, p, cumulative): returns the probability that a Binominal distributed random variable with parameters n (number of triels0 and p (success probability
  - = x (if cumulative = false)
  - ≤ x (if cumulative = true)

Similar to the Normal distribution function, the last parameter of the function cumulative is a boolean (with value true of false), indicating whether you wish to have the probability distribution function (true) or the probability density function (false). This holds for most probability density / distribution functions in MS-Excel.

MS-Excel contains many more standard probability distribution functions that you do not need for this course, but may be convenient for other applications. Examples are:

- POISSON.DIST(x, lambda, cumulative) : returns the probability that a Poisson distributed random variable with mean value lambda
  - = x (if cumulative = false)
  - ≤ x (if cumulative = true)
- EXPON.DIST(x, lambda, cumulative): returns the probability that an Exponential distributed random variable with rate lambda (i.e., mean 1/lambda)
  - = x (if cumulative = false)
  - ≤ x (if cumulative = true)

3.5 Causal models and MS Excel

Until here, we introduced how MS-Excel can help to *present* data. The real power of Excel is in supporting reasoning and decision-making, i.e., not just having primary data but the production of derivative data. If a spreadsheet should help in decision making and reasoning, we first need a good causal model that describes the decision problem. A causal model has variables, which contain data values, and their relations, which transform the values of an independent variable to a value for the dependent variable. In a spreadsheet both these data and transformational relations (formula) can be stored. These causal models may include "means-goals" statements, which are "if-then" or "what-if" statements helping to find the optimal set of means to achieve a specific goal, or knowing what the ambition can be given a limited set of resources. They also may be used to explain and predict, of course within a given set of assumptions that are part of the model and the values assumed for the variables included.

We first discuss how causal models can be made. MS Excel is not a drawing tool and consequently is not suitable for causal modeling. For this, we use MS-Visio by first selecting from the general templates 'Block Diagram'. This is a template by which one can represent values as blocks and causal relations as arrows. See Figure 3.33. (Note for readers who only want to learn using Excel here: They just can read the next few sections, without actually using MS-Visio). There are also oter free drawing tools like Lucidchart (www.lucidchart.com).

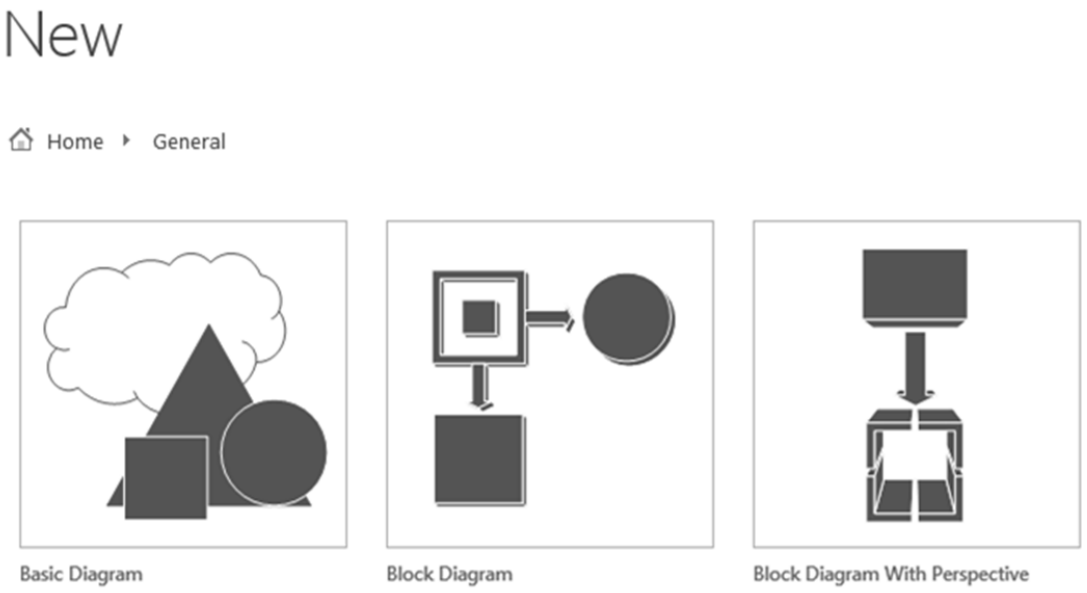


Figure 3.30. Possible templates that enable drawing causal diagrams in MS-Visio

This design contains data components and relationships. For Grand Cafe The Palace, profit is the main dependent variable, and costs have negative impacts on profit, whereas revenues have positive impacts on profit. There are three categories of fixed costs:

1. Electricity, water and gas
2. Management and administrative salaries
3. Housing (rent) and insurances.

There are several variable costs like:

1. Salary costs of cleaners, waiters and kitchen personnel
2. Purchasing costs of the drinks, meals and snacks.
3. These variable costs are highly dependent on the number of guests.

The revenues consist of

1. Drinks sold. Grand Palace has three categories here: cheap drinks, medium drinks and expensive drinks, with different sales prices and different purchasing costs.
2. Meals sold. Like the drinks, Grand Palace offers three categories here.
3. Snacks sold. Grand Palace offers different snacks, but all for the same price.

When you choose Block Diagram, you can use the different connectors and blocks to build a clear model. An example for Grand Café The Palace. See Figure 3.31.

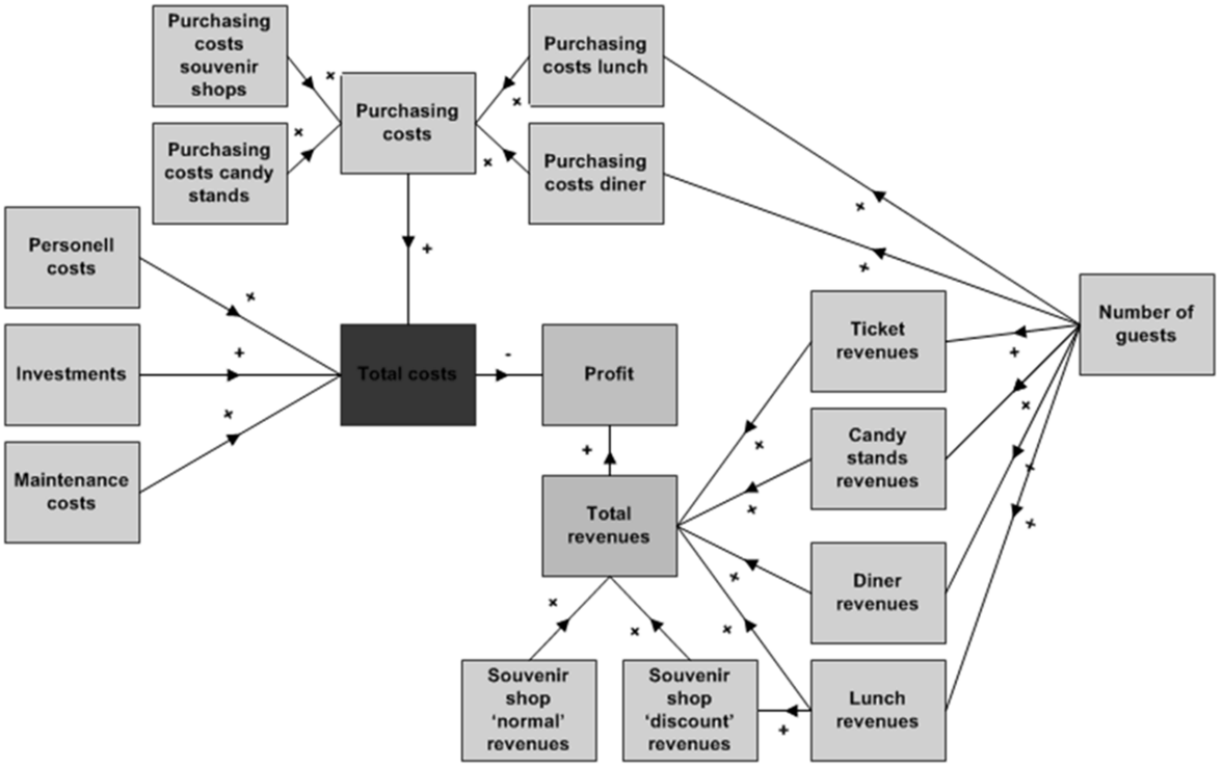
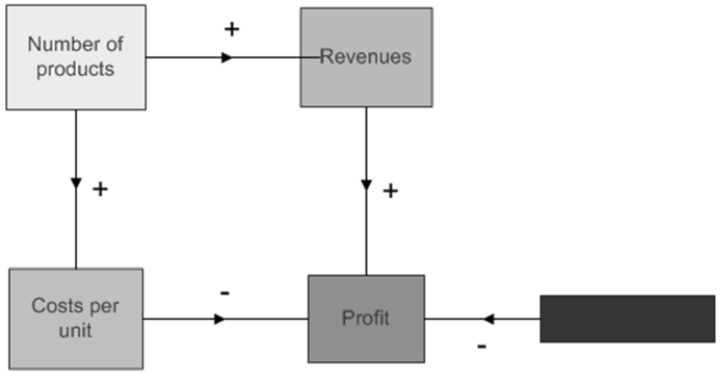


Figure 3.31. A causal diagram for Grand Café's costs and profit.

Note that a "+" stands for a positive relationship and a "–" for a negative relationship. A positive relationship means that when the box on the one side of the arrow becomes more, than the box on the other side of the arrow becomes also more or when the box on the one side of the arrow becomes less, than the box on the other side of the arrow becomes also less. A negative relationship means that when the box on the one side of the arrow becomes more, than the box on the other side of the arrow becomes less and the other way around.

Of course one can extend the model by adding a decision point or use the room left for comments under the model for pointing out some important aspects. Immediately you can see what kind of effect an input factor has on the final profit.

To see the link between the causal diagram and an MS-Excel spreadsheet, look at the following, simplified example in Figure 3.32, where each variable reappears as an item in the spreadsheet. Next formulas are given that express the causal relations between the (input and output) variables.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Fixed Costs	10												
2	Costs per unit	2												
3	Number of products	1	2	3	4	5	6	7	8	9	10			
4	Revenues	5	10	15	20	25	30	35	40	45	50			
5	Profit	-7	-4	-1	2	5	8	11	14	17	20			
6														
7	**COMMENTS:**													
8	The fixed costs of 10 are independent of the number of products													
9	they negatively influence the profit.													
10	The costs per unit are dependent of the number of products and they will rise													
11	in proportion to the rise of the number of products (so a positive relationship).													
12	The costs per unit influence the profit negatively (when the costs per unit rise,													
13	the profit will decrease).													
14	The number of products are given in this example, so this is a independent													
15	variable. The number of products positively influence the costs per unit and the													
16	revenues.													
17	The revenues are dependent of the number of products. This is a positive													
18	relationship. This will become clear when we look at the used formula													
19	(see the inserted comment).													
20	The revenues positively influence the profit.													
21	The profit is dependent on the costs per unit, the revenues and the fixed cost.													
22	These relationships are respectively negative, positive and negative. This													
23	will become clear when we look at the used formula (see the inserted formula).													

Gebruiker:
Formula used:
=$B$4*K3

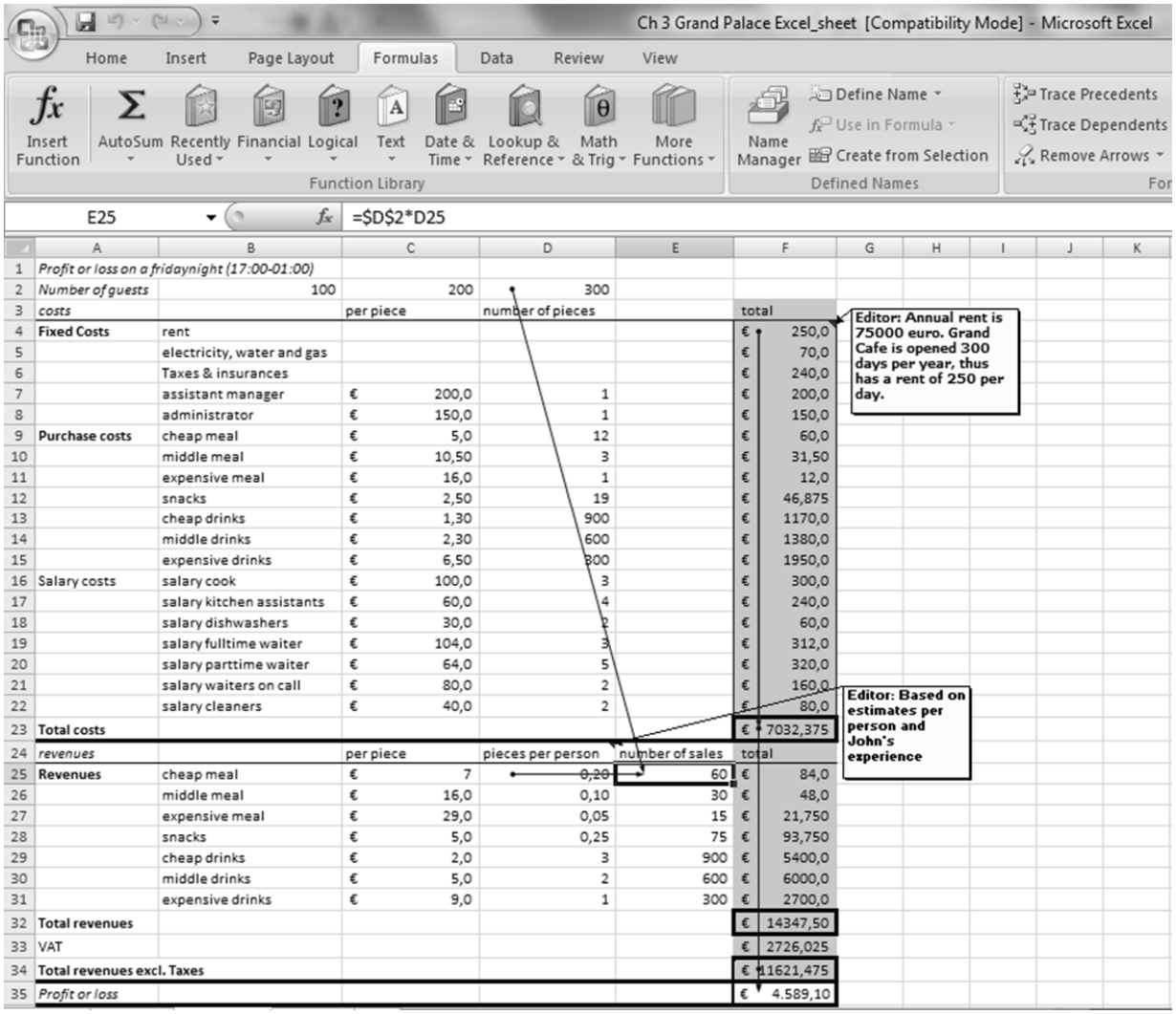
Gebruiker:
Formula used:
=K4-$B$1-K3*$B$2

Figure 3.32. Implementation of a causal diagram into a spreadsheet structure. Note the correspondence of colors between the variables.

Because the spreadsheet is made in the Dutch version of MS-Excel, “gebruiker” in the annotation fields is “user” in English.

If the models are well expressed in the spreadsheet structure, more advanced what-if analysis is possible to support reasoning and decision making. As stated in the beginning of this chapter, MS Excel (and spreadsheets in general) has excellent tools to help people making decisions on basis of causal understandings of relevant goals and means. A couple of examples are: How many employees do I have available for how many hours per month, and how much do they cost per month? What are the start-up costs of a certain machine and what is the payback time? How much can a project cost and what are realistic margins? What we like to see is what happens if something changes (e.g. the salary of the personnel). What is the impact of salary costs on the company’s profitability? Or how dependent is the organization on the number of visitors?

For our case, Grand Palace wants to analyze what would happen if more or less guests are present, or if people start consuming cheaper meals and drinks. To be able to do so, we need to present the causal relations mentioned in the beginning of this chapter in a spreadsheet design. This spreadsheet needs to present the fixed and variable cost parameters of the causal diagram, the revenue variables, the variable “number of guests” and we need to be able to define the strength of the causal relations between these variables. These variables and their parameters are given in Figure 3.33.



	A	B	C	D	E	F
1	*Profit or loss on a fridaynight (17:00-01:00)*					
2	*Number of guests*	100	200	300		
3	*costs*		per piece	number of pieces		total
4	**Fixed Costs**	rent				€ 250,0
5		electricity, water and gas				€ 70,0
6		Taxes & insurances				€ 240,0
7		assistant manager	€ 200,0	1		€ 200,0
8		administrator	€ 150,0	1		€ 150,0
9	**Purchase costs**	cheap meal	€ 5,0	12		€ 60,0
10		middle meal	€ 10,50	3		€ 31,50
11		expensive meal	€ 16,0	1		€ 12,0
12		snacks	€ 2,50	19		€ 46,875
13		cheap drinks	€ 1,30	900		€ 1170,0
14		middle drinks	€ 2,30	600		€ 1380,0
15		expensive drinks	€ 6,50	300		€ 1950,0
16	Salary costs	salary cook	€ 100,0	3		€ 300,0
17		salary kitchen assistants	€ 60,0	4		€ 240,0
18		salary dishwashers	€ 30,0	2		€ 60,0
19		salary fulltime waiter	€ 104,0	3		€ 312,0
20		salary parttime waiter	€ 64,0	5		€ 320,0
21		salary waiters on call	€ 80,0	2		€ 160,0
22		salary cleaners	€ 40,0	2		€ 80,0
23	**Total costs**					€ 7032,375
24	*revenues*		per piece	pieces per person	number of sales	total
25	**Revenues**	cheap meal	€ 7	0,20	60	€ 84,0
26		middle meal	€ 16,0	0,10	30	€ 48,0
27		expensive meal	€ 29,0	0,05	15	€ 21,750
28		snacks	€ 5,0	0,25	75	€ 93,750
29		cheap drinks	€ 2,0	3	900	€ 5400,0
30		middle drinks	€ 5,0	2	600	€ 6000,0
31		expensive drinks	€ 9,0	1	300	€ 2700,0
32	**Total revenues**					€ 14347,50
33	VAT					€ 2726,025
34	**Total revenues excl. Taxes**					€ 11621,475
35	*Profit or loss*					€ 4.589,10

Figure 3.33. A spreadsheet with "comments" and "trace precedents" for cell E25 (dependent variable "number of sales").

In Figure 3.33 we see the following….

- The blue arrows, called "trace precedents", show the causal relations in the spreadsheet. To see these arrows in MS Excel do the following:
  1. Click on a cell in which you entered a formula
  2. Click on the tab 'formulas'

3. Click on ‘Trace Precedents’. The arrows show where the output in the selected cell comes from/are based on. The other way around, if you want to show which cells are dependent of a specific cell, than:
4. Click on that specific cell
5. Click on ‘Trace Dependents’ (under the tab ‘formulas’)

- In cell E25, we reuse the parameter of the number of guest as given in cell D2. This cell has the absolute cell addressing sign $ before the column and row indicators ($D$2) so that the same number can be reused in E26 till E31 (if you copied the formula of cell E25 to E26 till E31). If we would not have done so, copying the formula from E25 to E26 would have change D2 to D3 etc. When one changes a selection of cells which contains formulas, the formulas also change. For example: when you have the formula ‘=SUM(B2:B10)’ in cell B11 and you move this to cell F51, the formula changes to ‘=SUM(F42:F50)’. This happens because the cell directions are connected with the position of the cell which holds the formula. In this case you must use absolute cell directions. You can do this by putting dollar signs before the coordinates (the row and the column indicators) in the original formula. That means that the formula from the example will be ‘=SUM($B$2:$B$10)’. When you copy this formula from cell B11 to cell F51 the formula will not change. But when you copy the formula ‘=SUM(B$2:B$10)’ to cell F51 (so, only dollar signs before the rows), the formula will change to ‘=SUM(F$2:F$10)’. Only the coordinate with the dollar sign in front of it, stays the same when copying the formula.
- As you see in the picture, one can insert comments for explaining your assumptions. You can use this for to explain why you have chosen a particular value or why a particular value is so high or low:
  1. Select the cell and press the right button of your mouse.
  2. Choose ‘Insert Comment’.
- In this spreadsheet, you calculate the profit or loss on an Friday evening. This particular evening is a normal Friday night. But what happens if it suddenly becomes very quiet in Grand Café The Palace; e.g. only 100 guests? For this, copy the whole sheet to another page, and change D2 into B2 for cell E25. You will also probably need less waiters, less cleaners, and less kitchen personnel.

To construct your spreadsheet in an easy way, it is convenient to copy formulas. To this end, you first have to select the cell or range of cells that you wish to copy, and next you can apply the standard method to copy and paste data (CTRL-C for copy, CTRL-V for paste). You can use SHIFT to select a contiguous

range of cells. Use CTRL-SHIFT-arrow key to select a column or row in the direction of the arrow key. Use CTRL to select several individual (non-contiguous) cells.

Relative versus Absolute cell references

By default, cell references in Excel are relative. See in Fig. 3.34 what happens if you use absolute or relative cell references while copying formulas.

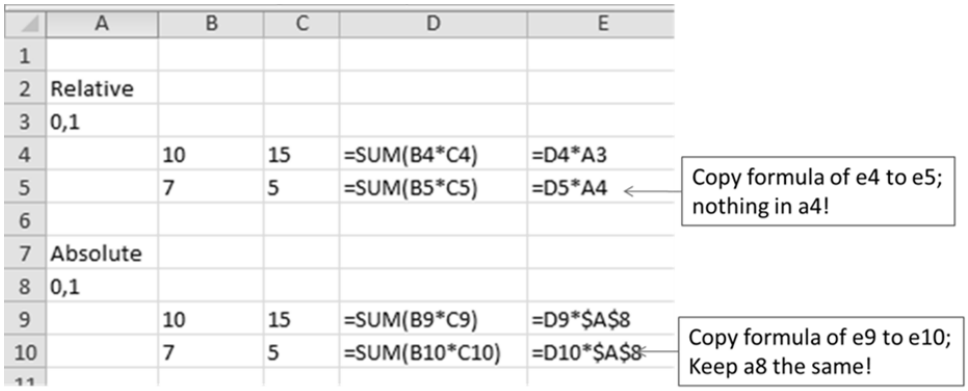


Fig. 3.34: Relative (default) versus absolute cell referencing

As you see, applying absolute cell references is very convenient if we want to always use the value of A8. To realize absolute cell referencing one needs to use a "$" before each letter and number that will be the reference value for the formulas (in our case $A$8). It is also possible to mix both relative and absolute cell references by referencing to A$8 for example, you would keep the row 8 as an absolute reference (it will not change), but the column A is relative and will change.

Another option of MS excel is to perform a "What If" analysis using a specific goal variable. For this we will use the "goal seek" function of Excel. Click on the Data tab, click on "What-if analysis" and there you will find the goal seek function.

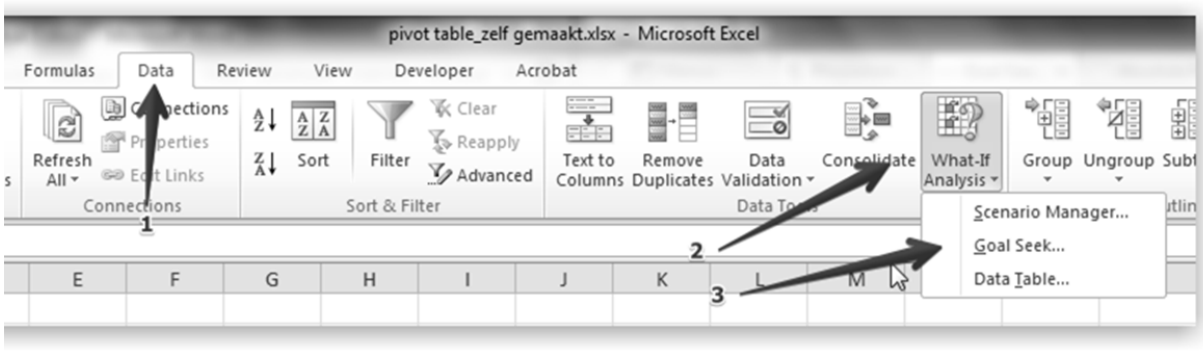


Fig. 3.35: Goals seek in the Data screen of Excel

For this example we have created a simple spreadsheet. The Goal Seek example is about a production process with a profit per item of 15 Euros. The total profit is determined by multiplying the number of items with the profit per item (=C3*C4). We want to know how many items we need to produce in order to get a total profit of 150 euros. Therefore we take as Set Cell the total profit cell (this is the cell Excel will monitor), we state that we want to achieve the value 150 and by changing the cell containing the number of items, which is C4. Click OK to continue See Fig. 3.36.

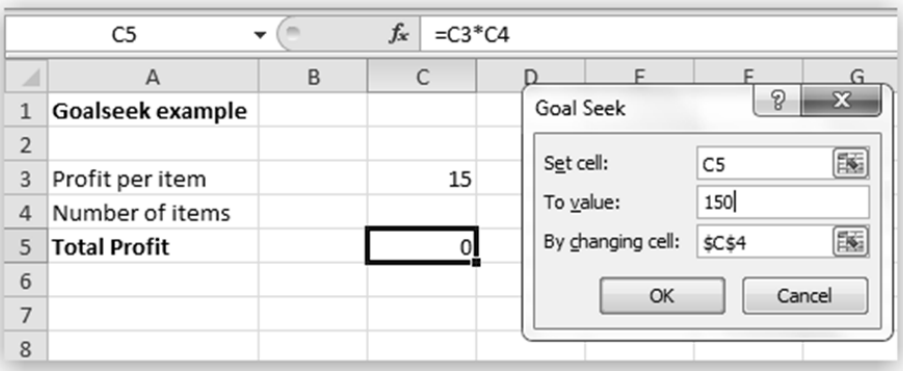


Fig 3.36: What number is needed to gain the value of 150 euro.

Goal Seek Status will appear showing the Target and Current value, along with that Excel shows the result in cell B3 (Multiplied by) that we required.

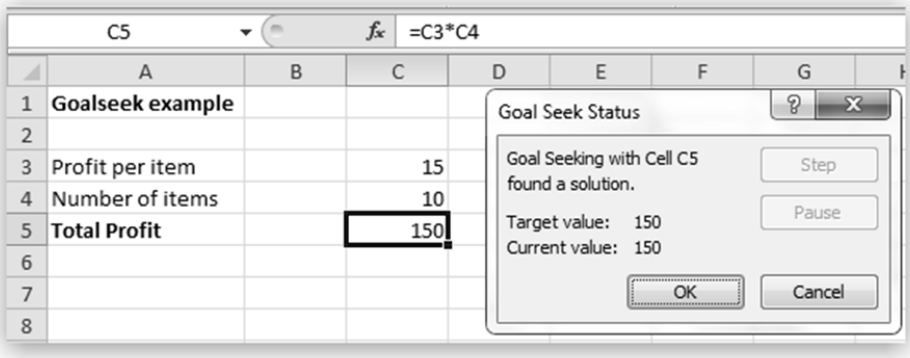
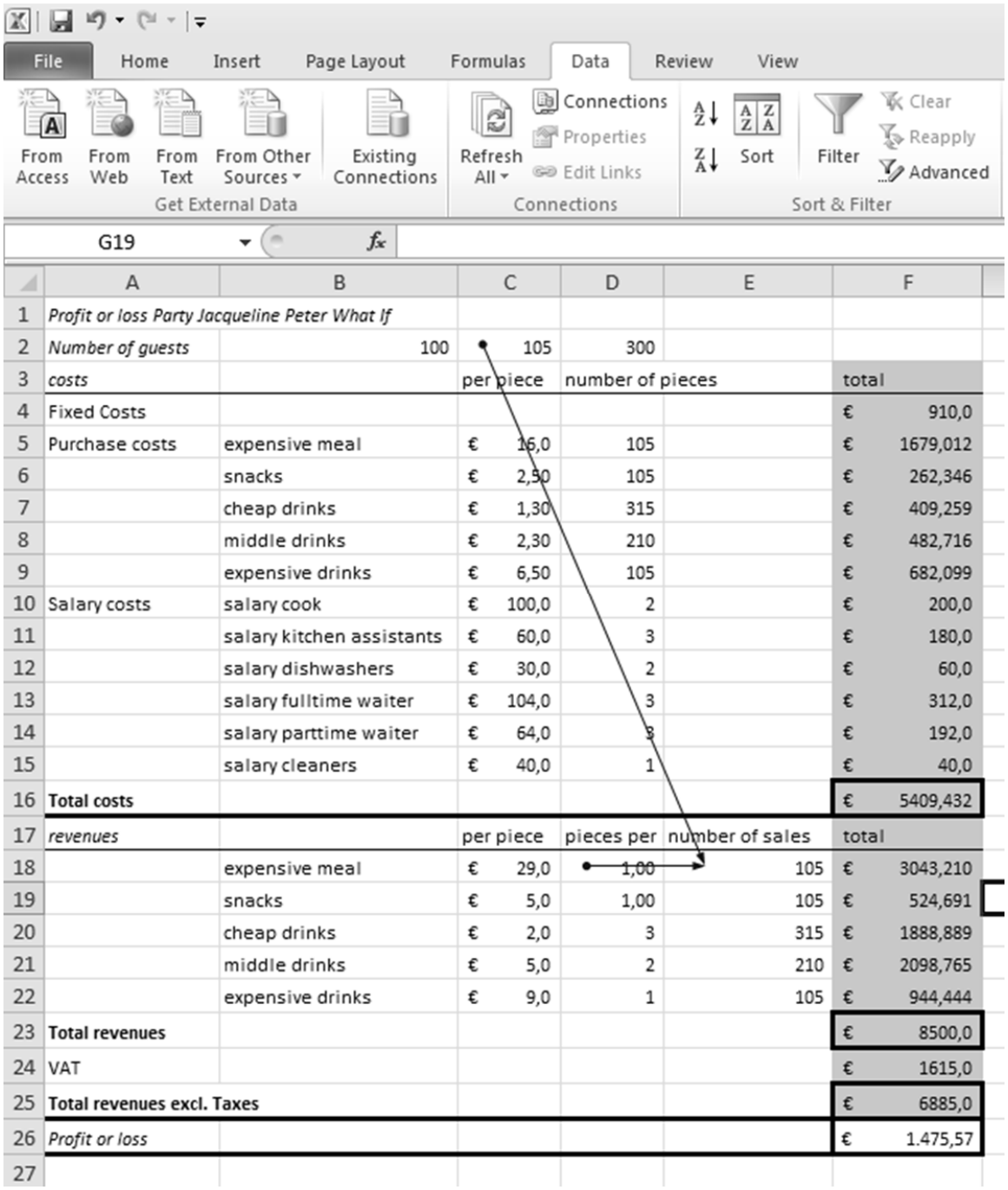


Fig 3.37: and the outcome of Fig 3.36

To continue our Grand Café case. Jacqueline and Peter want to hire the Grand Café to celebrate their 10th year of being together. They want to invite 200 guests, but do have only a budget of 8500 euros. So what are the possibilities? They want to treat their friends very well by offering a good meal. The price will be over 16,000, which is far above what Jacqueline and Peter can effort. John le Grand is willing to reduce the price by 1,000 euro, but still Jacqueline and Peter cannot make it. They do not want to reduce the `service level` to their visitors and consequently, they propose to reduce the number of visitors. For this, they performed a What-If-analysis (see Figure 3.38) by:

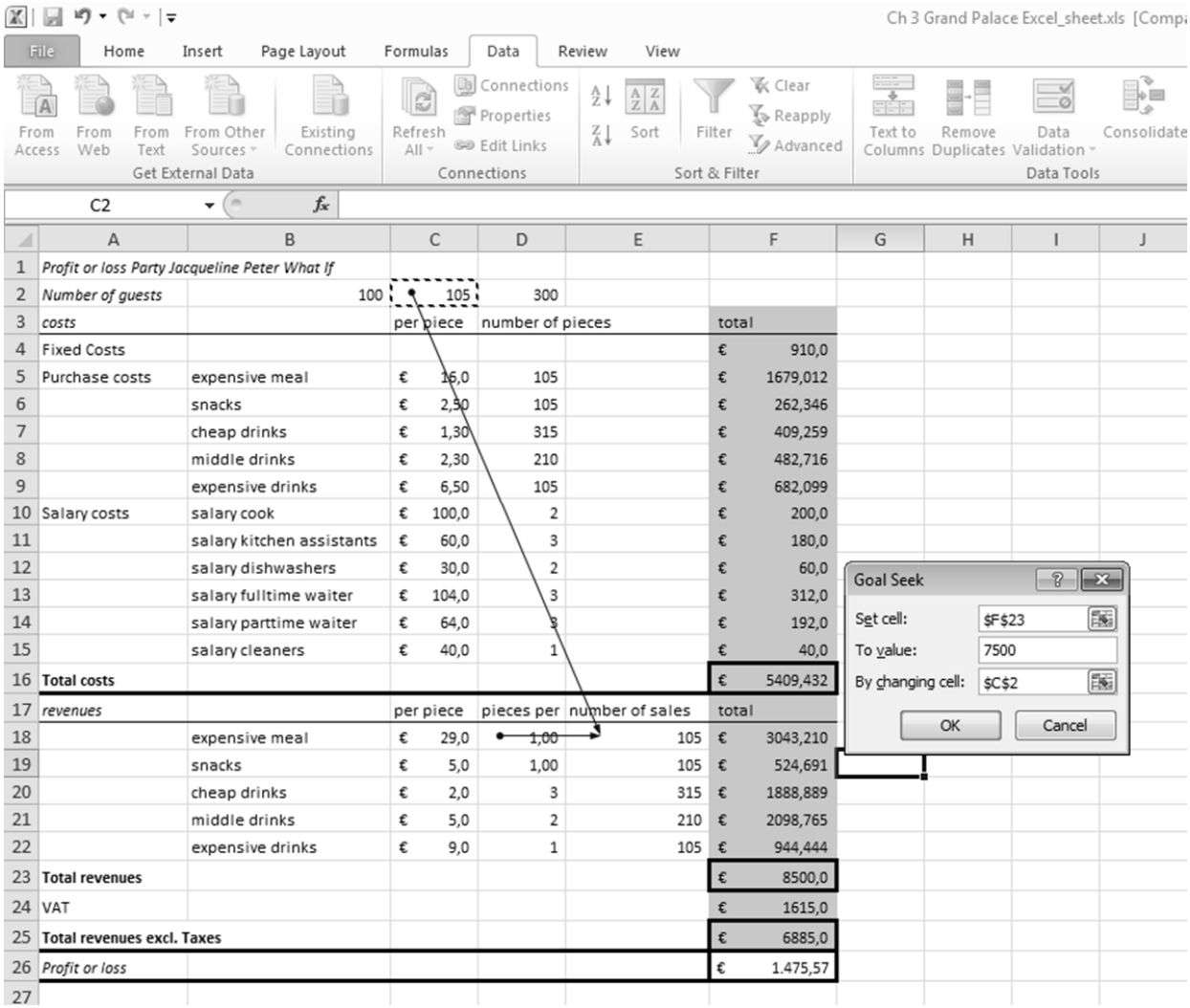
1. Click on 'Data'
2. Click on 'What-if-Analysis' and choose Goal seeking
3. Fill out the small screen in figure 3.37.

(Note you fill in 7500 instead of 8500, because they get 1000 euro discount.)



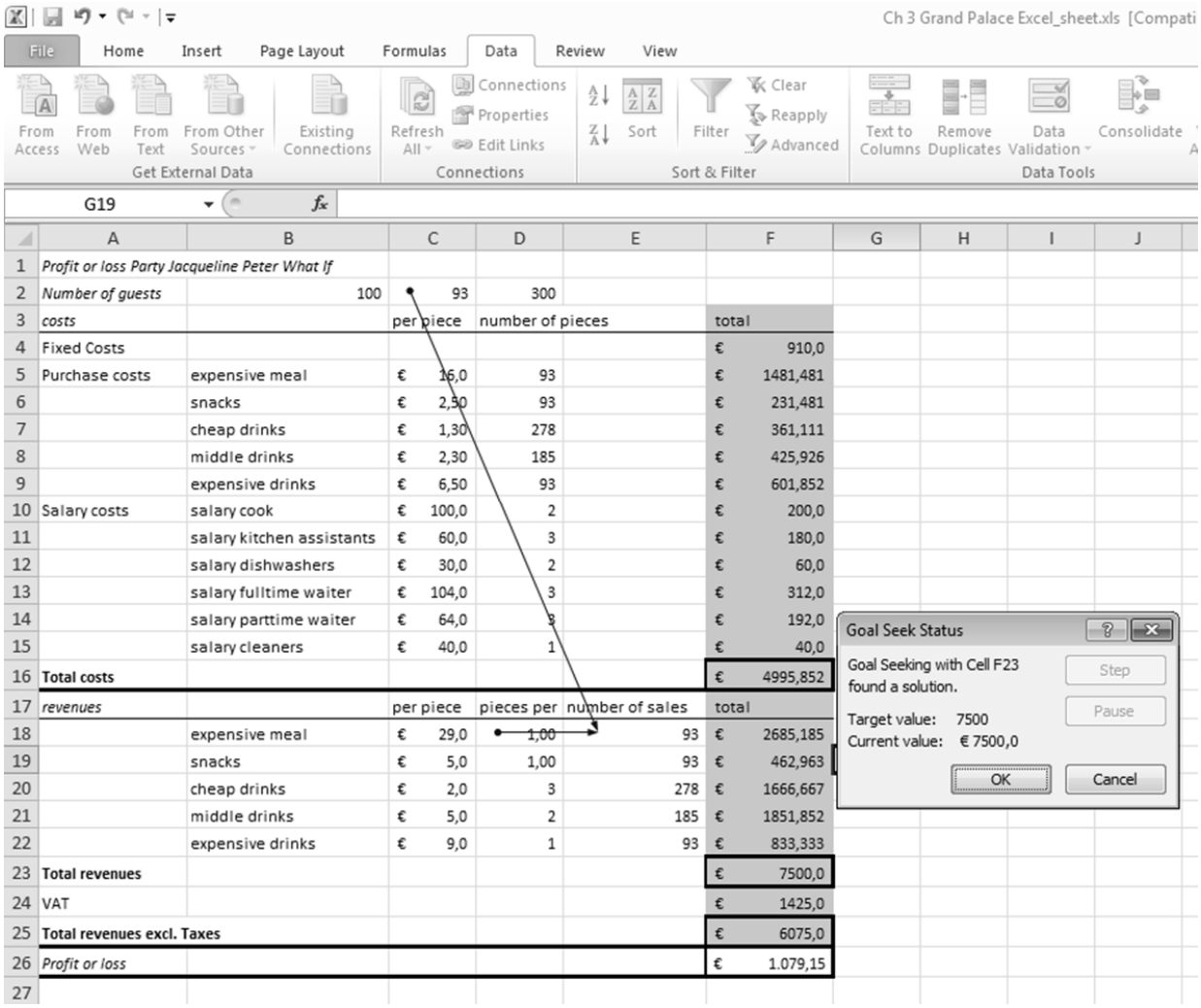
	A	B	C	D	E	F
1	*Profit or loss Party Jacqueline Peter What If*					
2	*Number of guests*	100	105	300		
3	*costs*		per piece	number of pieces		total
4	Fixed Costs					€ 910,0
5	Purchase costs	expensive meal	€ 16,0	105		€ 1679,012
6		snacks	€ 2,50	105		€ 262,346
7		cheap drinks	€ 1,30	315		€ 409,259
8		middle drinks	€ 2,30	210		€ 482,716
9		expensive drinks	€ 6,50	105		€ 682,099
10	Salary costs	salary cook	€ 100,0	2		€ 200,0
11		salary kitchen assistants	€ 60,0	3		€ 180,0
12		salary dishwashers	€ 30,0	2		€ 60,0
13		salary fulltime waiter	€ 104,0	3		€ 312,0
14		salary parttime waiter	€ 64,0	3		€ 192,0
15		salary cleaners	€ 40,0	1		€ 40,0
16	**Total costs**					€ 5409,432
17	*revenues*		per piece	pieces per	number of sales	total
18		expensive meal	€ 29,0	1,00	105	€ 3043,210
19		snacks	€ 5,0	1,00	105	€ 524,691
20		cheap drinks	€ 2,0	3	315	€ 1888,889
21		middle drinks	€ 5,0	2	210	€ 2098,765
22		expensive drinks	€ 9,0	1	105	€ 944,444
23	**Total revenues**					€ 8500,0
24	VAT					€ 1615,0
25	**Total revenues excl. Taxes**					€ 6885,0
26	*Profit or loss*					€ 1.475,57
27						

Fig. 3.38: The problem situation and causal model



	A	B	C	D	E	F	G	H	I	J
1	*Profit or loss Party Jacqueline Peter What If*									
2	*Number of guests*	100	105	300						
3	*costs*		per piece	number of pieces		total				
4	Fixed Costs					€ 910,0				
5	Purchase costs	expensive meal	€ 16,0	105		€ 1679,012				
6		snacks	€ 2,50	105		€ 262,346				
7		cheap drinks	€ 1,30	315		€ 409,259				
8		middle drinks	€ 2,30	210		€ 482,716				
9		expensive drinks	€ 6,50	105		€ 682,099				
10	Salary costs	salary cook	€ 100,0	2		€ 200,0				
11		salary kitchen assistants	€ 60,0	3		€ 180,0				
12		salary dishwashers	€ 30,0	2		€ 60,0				
13		salary fulltime waiter	€ 104,0	3		€ 312,0				
14		salary parttime waiter	€ 64,0	3		€ 192,0				
15		salary cleaners	€ 40,0	1		€ 40,0				
16	**Total costs**					€ 5409,432				
17	*revenues*		per piece	pieces per	number of sales	total				
18		expensive meal	€ 29,0	1,00	105	€ 3043,210				
19		snacks	€ 5,0	1,00	105	€ 524,691				
20		cheap drinks	€ 2,0	3	315	€ 1888,889				
21		middle drinks	€ 5,0	2	210	€ 2098,765				
22		expensive drinks	€ 9,0	1	105	€ 944,444				
23	**Total revenues**					€ 8500,0				
24	VAT					€ 1615,0				
25	**Total revenues excl. Taxes**					€ 6885,0				
26	*Profit or loss*					€ 1.475,57				
27										

Fig. 3.39: Using goal seek and setting the revenues at maximum 7500



	A	B	C	D	E	F
1	*Profit or loss Party Jacqueline Peter What If*					
2	*Number of guests*	100	93	300		
3	*costs*		per piece	number of pieces		total
4	Fixed Costs					€ 910,0
5	Purchase costs	expensive meal	€ 16,0	93		€ 1481,481
6		snacks	€ 2,50	93		€ 231,481
7		cheap drinks	€ 1,30	278		€ 361,111
8		middle drinks	€ 2,30	185		€ 425,926
9		expensive drinks	€ 6,50	93		€ 601,852
10	Salary costs	salary cook	€ 100,0	2		€ 200,0
11		salary kitchen assistants	€ 60,0	3		€ 180,0
12		salary dishwashers	€ 30,0	2		€ 60,0
13		salary fulltime waiter	€ 104,0	3		€ 312,0
14		salary parttime waiter	€ 64,0	3		€ 192,0
15		salary cleaners	€ 40,0	1		€ 40,0
16	**Total costs**					€ 4995,852
17	*revenues*		per piece	pieces per	number of sales	total
18		expensive meal	€ 29,0	1,00	93	€ 2685,185
19		snacks	€ 5,0	1,00	93	€ 462,963
20		cheap drinks	€ 2,0	3	278	€ 1666,667
21		middle drinks	€ 5,0	2	185	€ 1851,852
22		expensive drinks	€ 9,0	1	93	€ 833,333
23	**Total revenues**					€ 7500,0
24	VAT					€ 1425,0
25	**Total revenues excl. Taxes**					€ 6075,0
26	*Profit or loss*					€ 1.079,15
27						

Figure 3.40: And the result: reduce number of guests to maximum 93

3.6 Causal models and uncertainty

The Data Table function is a function that allows the user to quickly pose a range of "what if" questions. It is very useful for sensitivity, variance analysis and Monte Carlo simulation. You can find the Data Table function under the "Data" tab by clicking on the "What-if analysis" button (see Fig. 3.41).

For demonstration of the Data Table function we will continue with the example of the Grand Café. For this part we simplify the situation somewhat in order to maintain the focus on the way these calculations are performed. Below, in figure 3.41, we show the causal model of the Grand Café's profitability. In this simplified version the revenue, the personnel cost and the costs for food and drinks are dependent on the

number of costumers on any given day. In figure 3.41 you can see the profit calculation. In this calculation we make the following assumptions:

1. The grand café needs a minimum number of 5 employees to make sure that they can help their costumers
2. If there are more than 50 customers, per group of 20 customers on top of this an extra employee is needed. So for 90 customers 5 + 2 = 7 staff members are needed.
3. Costs of food and drinks are only dependent on the number of costumers. We do not take storage nor spoilage into account.

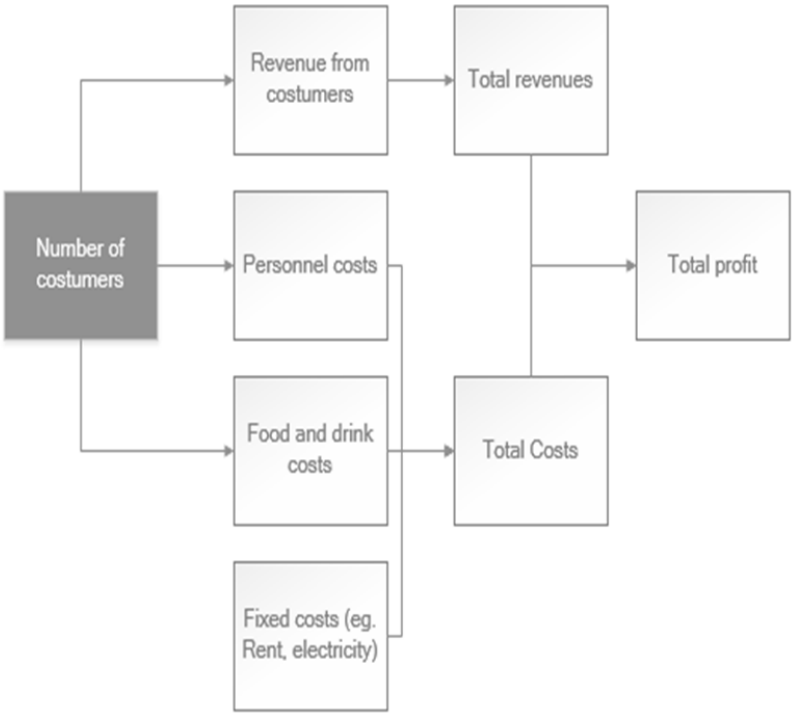


Figure 3.41: Causal model of Grand Café's profitability

	A	B	C
1			
2		**Formula**	**Result**
3	Number of Costumers	200	200
4	Personnel needed	=IF(B3>50;5+ROUNDUP((B3-50)/20;0);5)	13
5			
6	Revenue from costumers	=B3*35	7000
7			
8			
9	Costs		
10	Food and drink	=B3*10	2000
11	Personnel	=B4*100	1300
12	Fixed	300	300
13			
14			
15	Profit	=B6-SUM(B10:B12)	3400
16			

Fig. 3.42: Excel Model

In Figure 3.42 you can see that we first fill in the sheet to calculate the profit of the Grand Café. As a reference you can see the formulas in the column titled “Formula”. Of course you do not need to show these formulas, when you type these in, you will only get the calculated result, as shown in the column “Result”.

In this example we can see that if there are 200 costumers, all of whom spend €35,- on average, the Grand Café makes a profit of €3.400,- But what if the average spending is lower, or the number of costumers is different. This is what the Data Table function is made for. Next to the model add a column “Number of costumers” and a column “Profit”. In the column number of customers fill in a range of different possible customer numbers. The column profit you should leave empty, as we will fill this column automatically with the Data Table function. However, make sure to leave the first row of the number of customers empty and the row of Profit should be “=B15”, so that it refers to the cell containing the profits (see Figure 3.43).

	A	B	C	D	E
1					
2		**Result**		**Number of costumers**	**Profit**
3	Number of Costumers	200	People		=B15
4	Personnel needed	13	Personnel	50	
5				100	
6	Revenue from costumers	7000	Euros	150	
7				200	
8				250	
9	Costs				
10	Food and drink	2000	Euros		
11	Personnel	1300	Euros		
12	Fixed	300	Euros		
13					
14					
15	Profit	3400	Euros		
16					

Leave this Empty

Make sure to reference B15

Fig 3.43: preparation for Data Table

The next step is to build the data table. It is very important to first select the cells containing the number of customers and the empty profit cells. So in this case select the cells from D3 to E8, go to the tab “data”, the What-if button and then select Data Table.

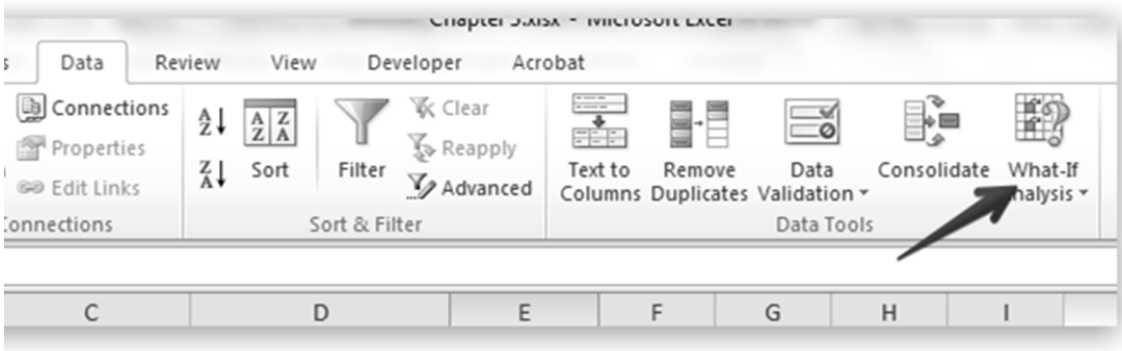


Fig 3.44: What-if analysis button

After clicking the Data Table option, click on the button next to the "Column input cell" and select the cell containing the number of customers. What this does is that it tells MS-Excel to use the column with the number of customers you created and fill this in the cell B3 in order to calculate the profits for each number. Click OK and the data table should fill up, you can now see what the profits would be for different numbers of customers. See the picture below for help.

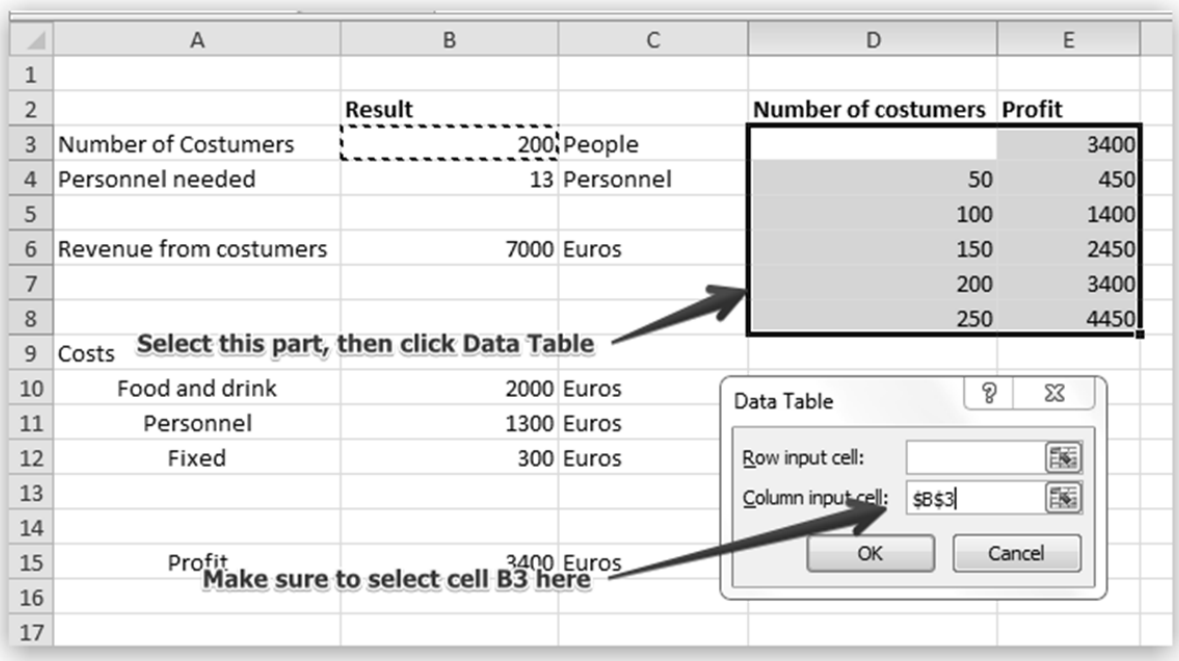


Fig 3.45: Data Table

When we plot the data from our Data Table in a graph, one can see that the profits increase in a linear fashion when with the number of customers increasing.

Fig 3.46: Graph of profits

Note: had our input data been arranged horizontally in Rows, we could have used a Row Input Cell to process the data instead of the Column Input Cell we used in the example.

Another Scenario is often where you want to vary an input by a Fixed Percentage. This is easily done using Data Tables.

- Setup the input cells with the percentage variations you want to examine, noting that the values don't have to be evenly spread.
- Setup a Temporary Input Cell, This will hold the Percentage Variance briefly whilst calculations are happening. Set a default value of 0 (zero)
- Change your Main Input Cell, customers in our case, to multiply the fixed answer by 1+ the temp Input Cell.
- Run the Data Table with a Column Input Cell, which will refer to the Temp Input Cell.

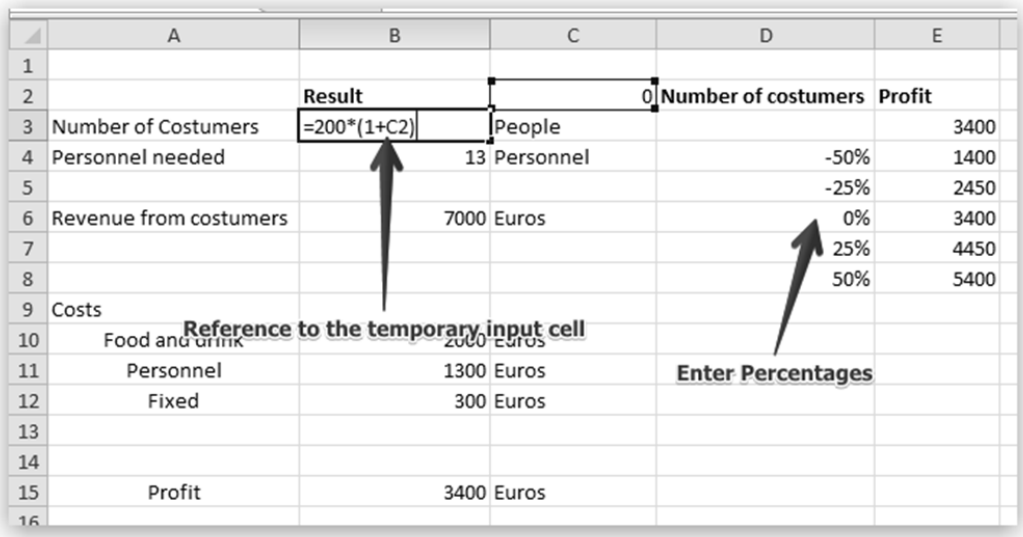


Fig. 3.47: Applying percentages of changes

But what if not only the percentage of expected customers changes, but also the fixed costs are different? We can use a Data Table with both the Row and Columns as input to generate a matrix of the profits. Use the column with the percentages from the previous example. Make a new row with Fixed costs and use the cell in the corner (D3) to reference to the profit, as this is what you want to monitor.

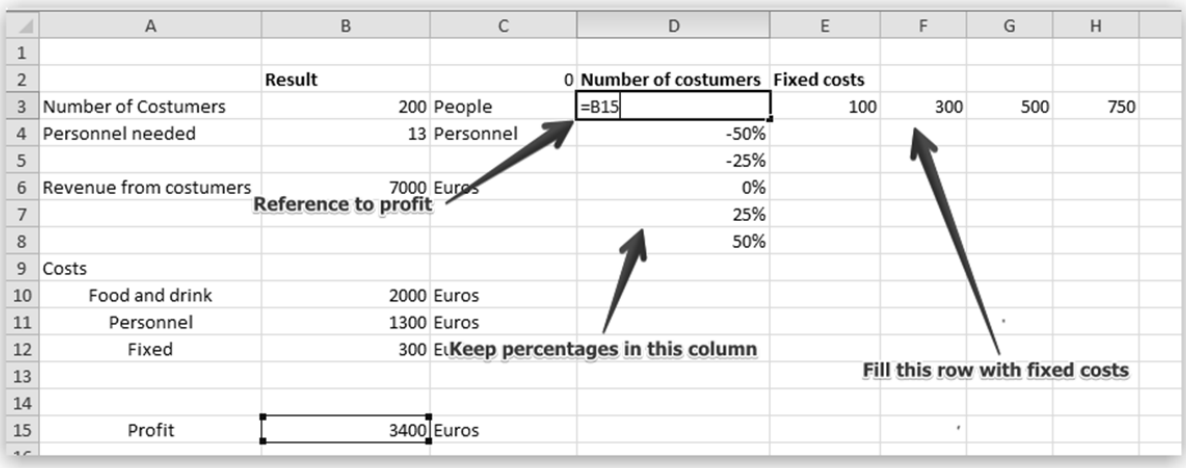


Figure 3.48: Generating a matrix of profits outcome for two varying independent variables

Next step, select all the cells for the Data Table as shown in the picture below. Open the data table menu, for the Row input cell (this contains the fixed costs row) select the cell of the fixed costs. For the Column input cell select the temporary cell C2. Press OK and the data table should fill up. Now you can see for example what the profits are with 750 euros fixed costs and 50% more customers than the expected 200 (4950 euro profit in this case).

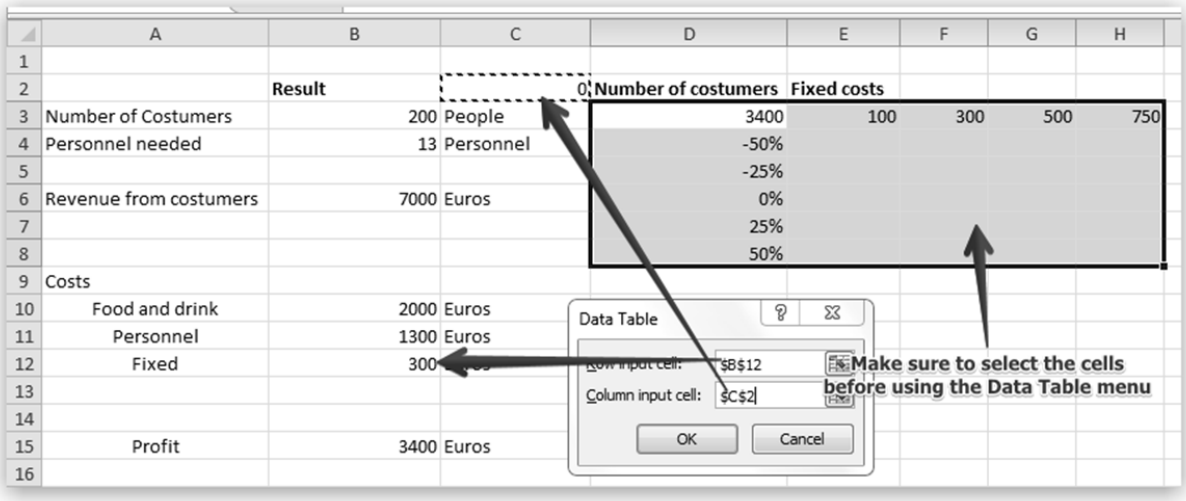


Figure 3.49: Output of Figure 3.48

What if you have a complex causal model and want to monitor a number of independent and dependent variables at once? In this example we are varying one independent input variable by monitoring 2 dependent output variables, namely profit and revenue. What you will do is create three columns, which labels are “Costumers”, “Profits” and “Revenue”. The column costumers is filled as before (leaving the first row empty, the column “Profits” has a reference to B15, the cell containing the profits in the first row. Similarly the column Revenue refers to the cell containing the revenue (B6).

The next step is the same as explained before, select all three columns, open the Data Table menu. In this menu select the temporary input cell C3 as column input cell. Click ok and both the profit and the revenue for the numbers of costumers will be filled in (see Figure 3.50).

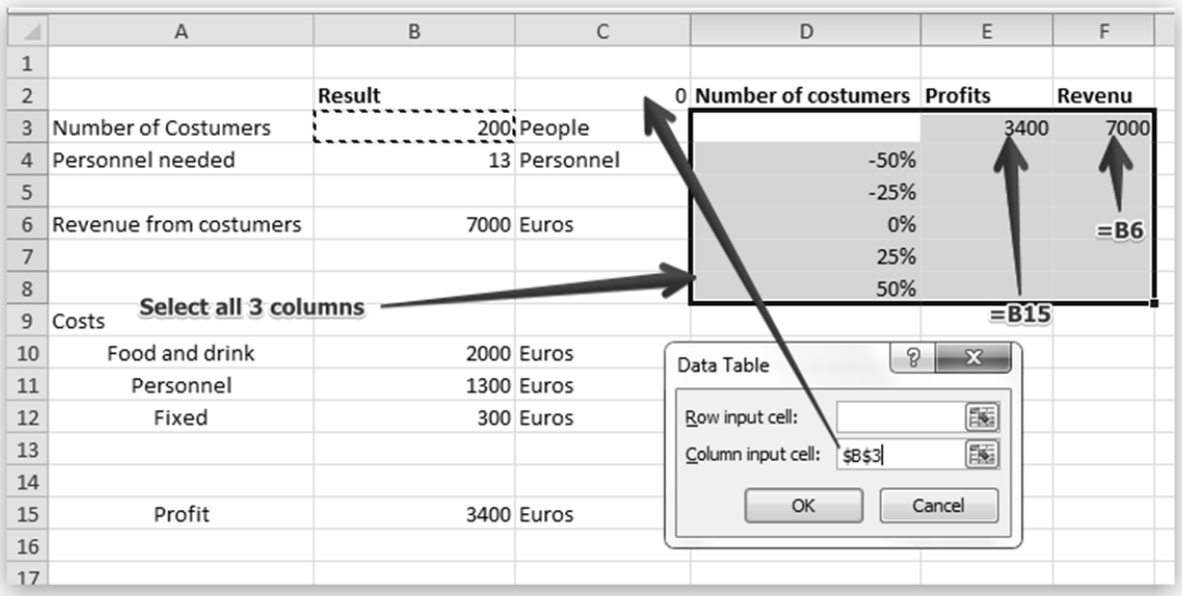


Fig. 3.50: Monitoring multiple variables

The data table function has many applications. Another instructive example of it is given by Hui in his www.chandoo.org site (accessed September 15, 2013) in which he illustrates Monte Carlo simulations using MS Excel by the case of a gold mine. See http://chandoo.org/wp/2010/05/06/data-tables-monte-carlo-simulations-in-excel-a-comprehensive-guide/.

Monte Carlo simulation allows the user to measure the effect of variability of the input parameters. This works by changing the input variables hundreds or thousands of times to simulate variability and measure the effects at the end of the runs.

The variability which the Monte Carlo simulation simulates is not always the same however. This is because variability is not the same for all processes. Most processes are variable around an average value, the spread around this average value is called the distribution. There are many different distributions, but the three most common ones are:

1. Normal: Bell shaped around a mean
2. Uniform: All values have an even chance of selection
3. Exponential: Low or High values have a much higher probability that the other values

Most distributions are Normal distributions, therefore in our example of Monte Carlo simulation we use the Normal distribution to calculate variability. However this is not necessarily the case! Make sure that you use a distribution that is appropriate for your model.

There are two functions in MS-excel that can generate random numbers: Rand and Randbetween. Both these functions have a uniform distribution however, meaning that all the values between the minimum and maximum values will have an equal probability of being generated. Since we want to work with a Normal distribution we will convert these numbers with the "norminv" function in the following as follows: "=norminv(rand(),mean,standard_dev)". The first argument rand() generates a random number between 0 and 1, for mean we will enter a mean value for our process and standard_dev is the standard deviation (the spread of the variability) we want our process to have.

For example the function "=norminv(rand(), 200, 10)" generates a distribution of random numbers centered on 200 with a spread having a bell shaped curve with a standard deviation of 10. This means that the function will produce a number with a 99.7% probability of being between 170 and 230 and on average will have a mean of 200.

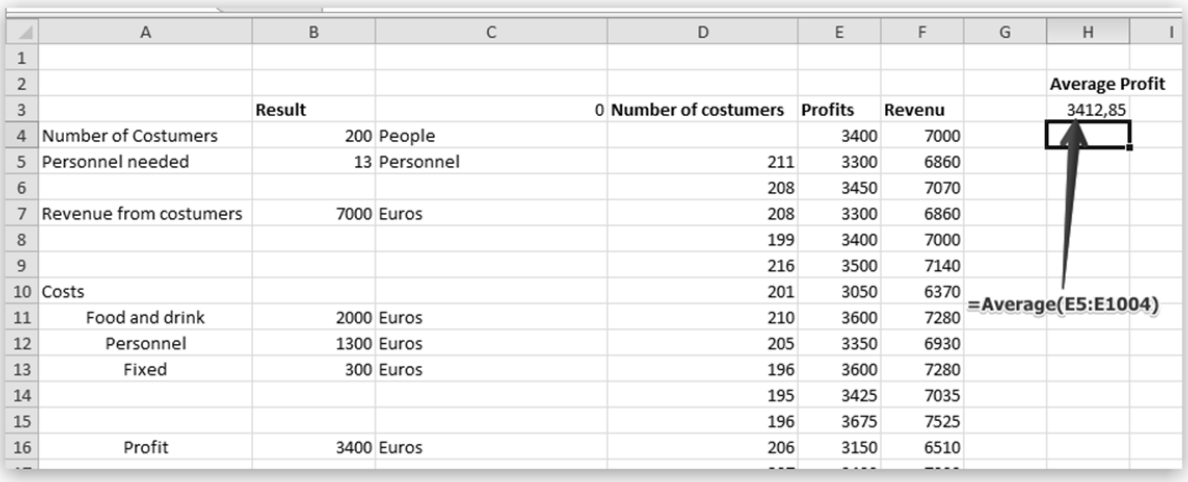
How to use this knowledge combined with Data Tables to do Monte Carlo simulations?

SUM =ROUNDUP(NORM.INV(RAND();200;10);0)

	A	B	C	D	E	F	G
1							
2							
3		**Result**	0	**Number of costumers**	**Profits**	**Revenu**	
4	Number of Costumers	200	People		3400	7000	
5	Personnel needed	13	Personnel	=ROUNDUP(NORM.INV(RAND();200;10);0)			
6				185			
7	Revenue from costumers	7000	Euros	201			
8				209			
9				197			
10	Costs			193			
11	Food and drink	2000	Euros	189			
12	Personnel	1300	Euros	199			
13	Fixed	300	Euros	206			
14				210			
15				197			
16	Profit	3400	Euros	185			

Fig. 3.51: Applying a normal distribution in a Monte Carlo simulation

As stated above the formula "=NORMINV(RAND(),200,10)" gives an average of 200 customers with a standard deviation of 10. However since customers are integer (you cannot have 200.8 customers) we will round the generated number up to get a rounded figure. Therefore the final formula we use is "=ROUNDUP(NORM.INV(RAND();200;10);0)". Fill in this formula for 1000 rows (select the cell and drag it down) (see Figure 3.51). Then select all three columns for 1000 rows and use the steps as described before to generate a Data Table and you end up with the content of Figure 3.52.



	A	B	C	D	E	F	G	H	I
1									
2								Average Profit	
3		Result	0	Number of costumers	Profits	Revenu		3412,85	
4	Number of Costumers	200	People		3400	7000			
5	Personnel needed	13	Personnel	211	3300	6860			
6				208	3450	7070			
7	Revenue from costumers	7000	Euros	208	3300	6860			
8				199	3400	7000			
9				216	3500	7140			
10	Costs			201	3050	6370			
11	Food and drink	2000	Euros	210	3600	7280			
12	Personnel	1300	Euros	205	3350	6930			
13	Fixed	300	Euros	196	3600	7280			
14				195	3425	7035			
15				196	3675	7525			
16	Profit	3400	Euros	206	3150	6510			

Fig. 3.52: Monte Carlo simulation with input variable Number of customers between 170 and 230 with a mean of 200 allowing 1,000 iterations between E5:E1004.

In the model above you can see that while the estimated profit was 3400 euros, using 1000 iterations of Monte Carlo simulation we actually have an average profit over all these trials of 3412,85 Euro, which means that the business model is very robust against any kind of variations in values of the main independent variables involved on the long run of 1,000 iterations.

Note 1: The used formula is a Volatile Formula, i.e, it recalculates every time the worksheet changes. What this means for the worksheet is that when the Data Table goes to Calculate Row 2 of the Data Table it will recalculate the Input value for Row 1. On Calculation of Row 2, it doesn't change the Table Values for Row 1, just the Input Column value.

Note 2: Always run at least 1000 iterations of Monte Carlo models. This is to ensure that you have a statistical chance of getting sufficient outliers (extreme values) to make the variance analysis meaningful.

This is important because as the number of iterations increases the variance of the average output decreases.

Try pressing F9 a few times and watch the average change (wait a second or two to allow your computer to perform the calculations).

3.7 MS-Excel solver add-in, analysis pack, regression and macros

The Solver is an add-in for MS-Excel which is used for the optimization and simulation of business and engineering models. It solves complex linear and nonlinear problems and can also be used in conjunction with VBA to automate tasks. The Solver is included by default in MS-Excel but kept disabled. In order to enable it, click the File Menu and choose Options.

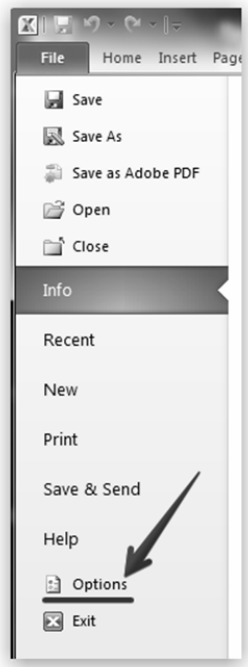


Fig 3.53 Select Solver add-in

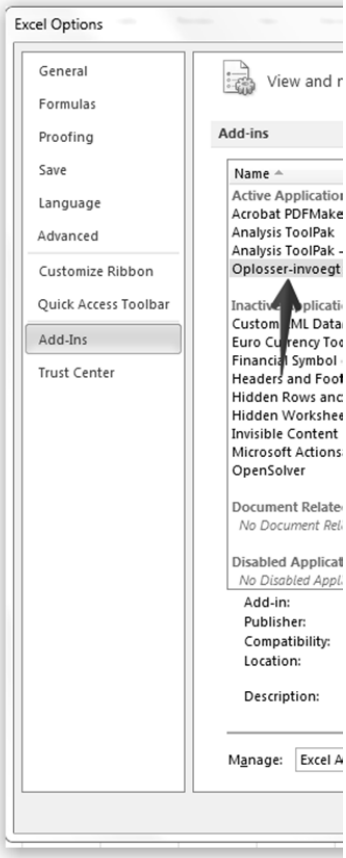


Fig. 3.54: Enabling the solver add-in

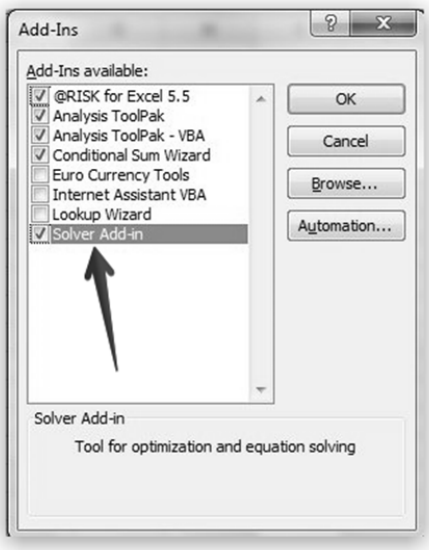


Fig 3.55: Solver located under Data tab

In the options menu click on the Add-Ins tab. Select the Solver Add-in and click the Go button at the bottom.

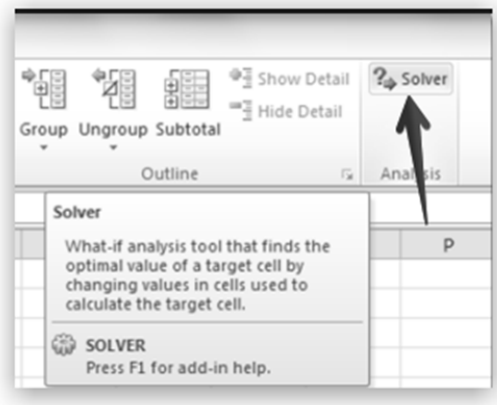


Fig 3.56 Solver enabled

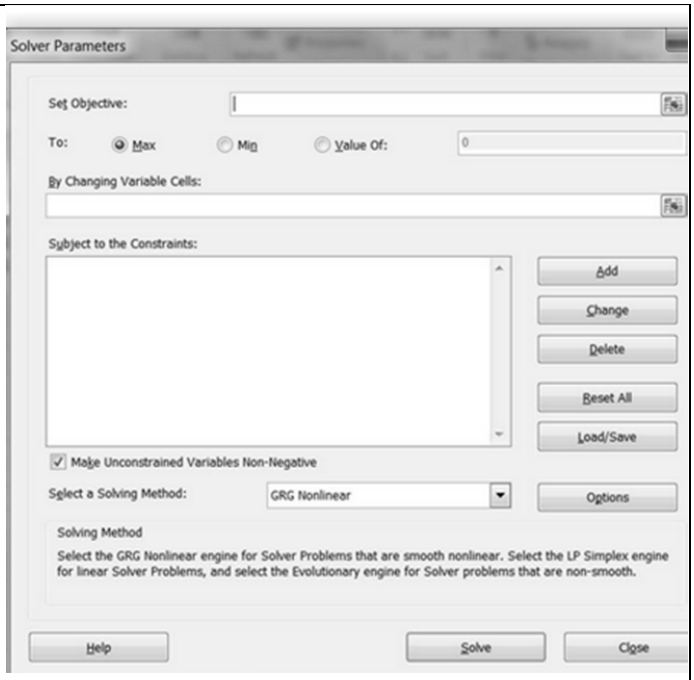


Fig 3.57: Solver dialogue box

When you click the Solver button, the solver screen is displayed as shown in Fig. 3.57. This screen is used to provide the parameters for the solver engine. The parameters that you provide are dependent on the problem that you want to solve, but we will give an overview of the workings of the solver module.

First start with setting the objective, this is the target cell that you want to optimize. So for example you would set the Profit as the target cell in order to determine the way to maximize profits. Or you could use the minimization option and set the Costs as target cell, this way determining how to minimize the costs of your operation.

Secondly you enter the variable cells that need to be changed. These are the cells that determine the profit and the solver module will try to find the configuration of the different variables that maximizes or minimizes the outcome.

Finally, you can add constraints for the solver engine. This is used to make sure that the solver only can use the amount of resources you have and does not just use enormous amounts of money/time/personnel in order to maximize the profits.

Once you have configured everything, like in Figure 3.58 click to the solve button to let the Solver try to solve your problem.

Note: Many more add-ins are available on internet as freeware.

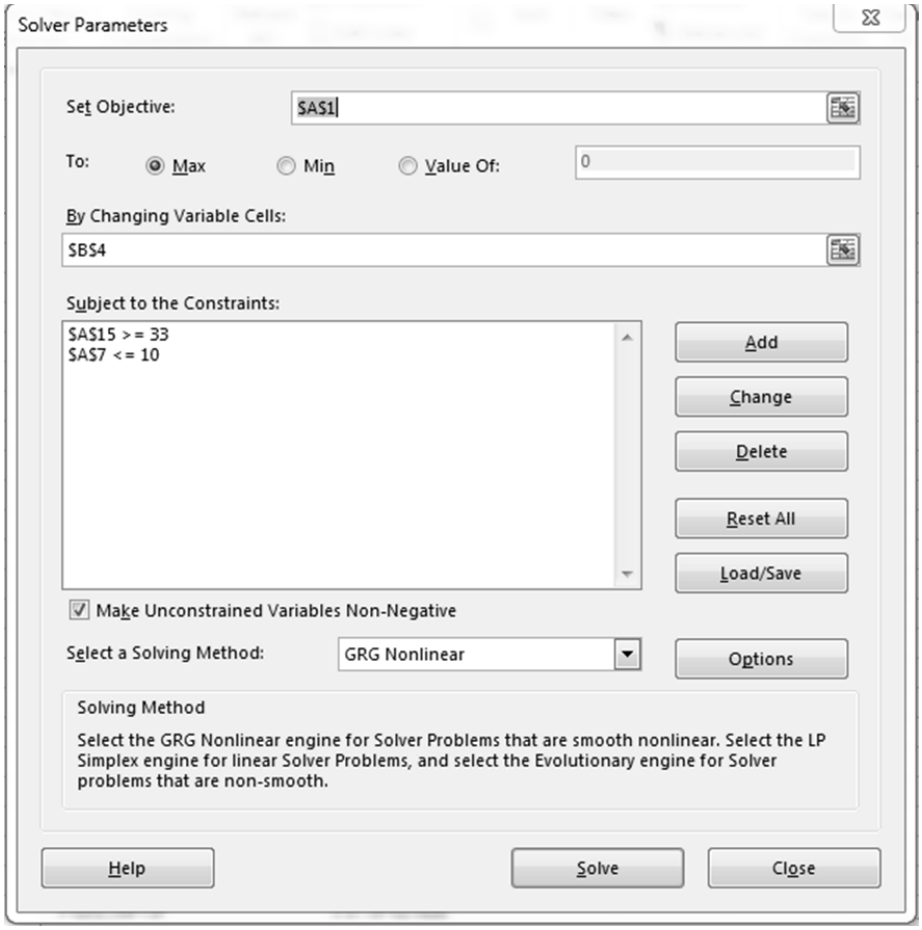


Fig. 3.58: Solver parameters

Data Analysis

Next to the many built-in functions of MS-Excel there are many additional statistical functions which are provided by the analysis tool pack. Just like the Solver, you must install this module first before you can use it. Once installed, go to the data tab and click the Data Analysis button.

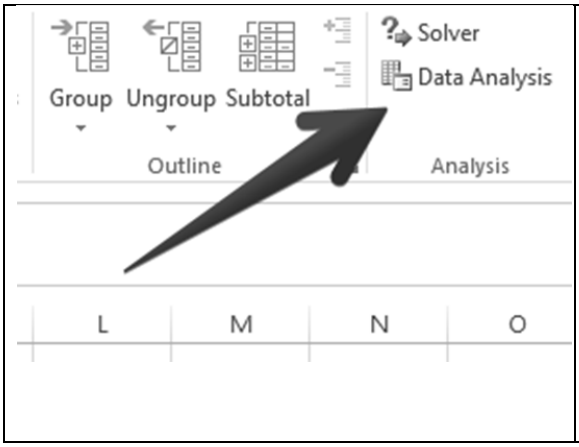


Fig. 3.59. Data analysis option

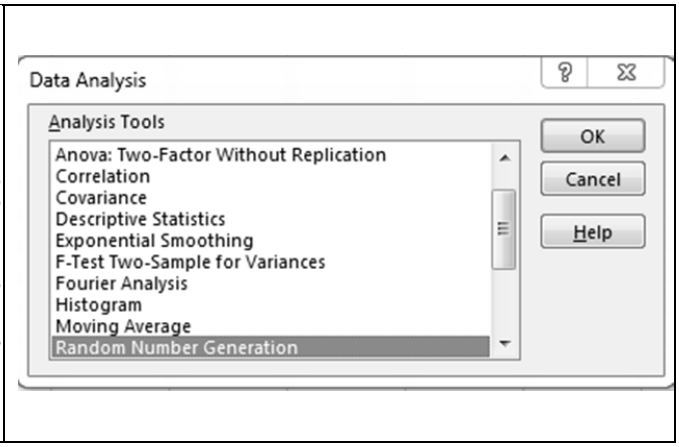


Fig. 3.60. Data analysis selection

From the many functions available we will focus on the histogram and regression function. Suppose you want to make a histogram of the following values ("Input range"), with the following bins ("Bin range").

	A	B
1	**Input range**	**Bin range**
2	87	20
3	27	40
4	45	60
5	62	80
6	3	
7	52	
8	20	
9	43	
10	74	
11	61	

Fig. 3.61 Input data and Bin range

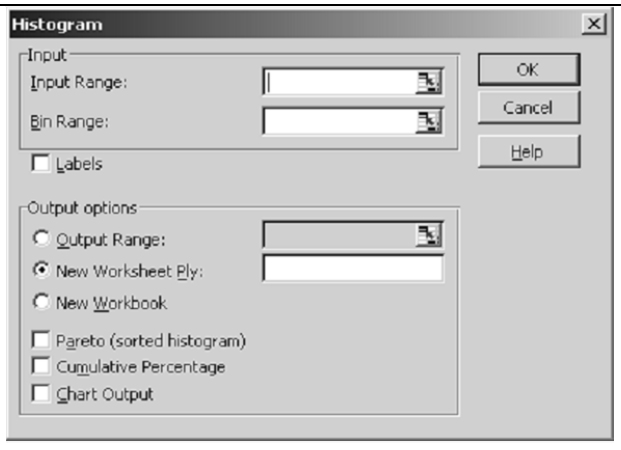


Fig. 3.62 Histogram selection box

You can do this with the Analysis ToolPack, where you have to select Histogram. The following window appears (see Fig. 3.63). In the Input Range box, type A2:A10; in the Bin Range box, type B2:B4 (in both cases we do not include the headers, resp. A1 and B1). Under Output Options, click New Workbook, select the Chart Output check box, and then click OK.

A new workbook with a Histogram table and an embedded chart is generated. MS-Excel counts the number of data points in each data bin. A data point is included in a particular data bin if the number is greater than the lowest bound and equal to or less than the greater bound for the data bin. In the example here, the bin that corresponds to data values from 0 to 20 contains two data points, 3 and 20. If you omit the bin range, MS-Excel creates a set of evenly distributed bins between the data's minimum and maximum values.

The information for this example is taken from the Microsoft Support website (http://support.microsoft.com/kb/141684).

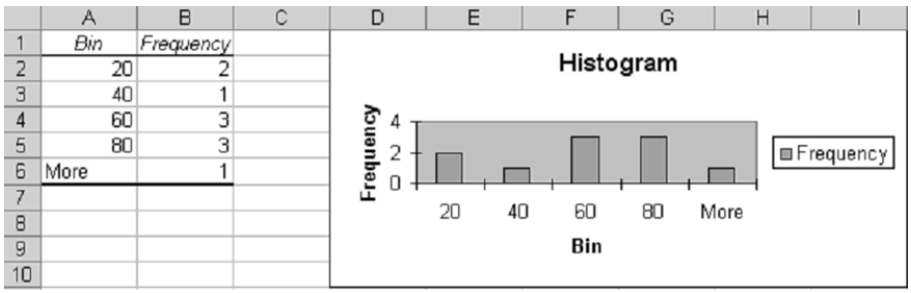


Fig. 3.63: Output of analysis toolpack for Figures 3.61 and 3.62 data.

The Correlation tool creates a table of correlation coefficients for all possible pairs of 2 or more independent variables in a range. It looks as follows:

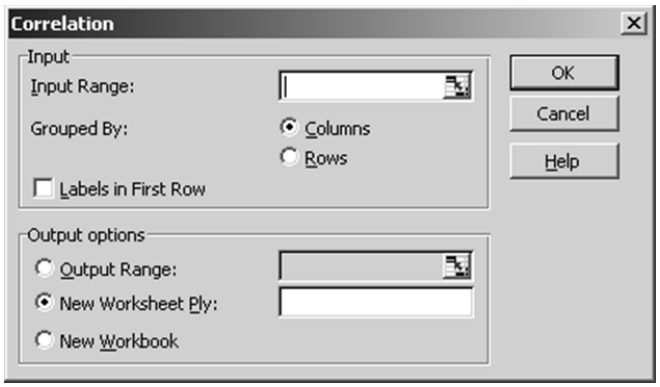


Fig. 3.64: Correlation dialogue box

You can simply enter the data range and indicate whether the data are grouped in rows or columns. The output will be a matrix of all pairwise correlations. Regression analysis is used for analyzing and modeling variables, looking at the relationship between independent variables and a dependent variable. In simple terms, a regression analysis shows how the dependent variable changes when the independent variables change. Therefore this shows the causality between variables, but with relational probabilities.

You can access the regression function from the Data Analysis menu. Selecting the regression analysis option will launch the dialogue box which will let you apply the regression feature on your data. The regression analysis screen is shown in Fig. 3.65. The input Y range is for selecting the dependent variable, the input X range is for selecting the independent variable. Pressing Ok will generate the regression outputs (see Figure 3.66).

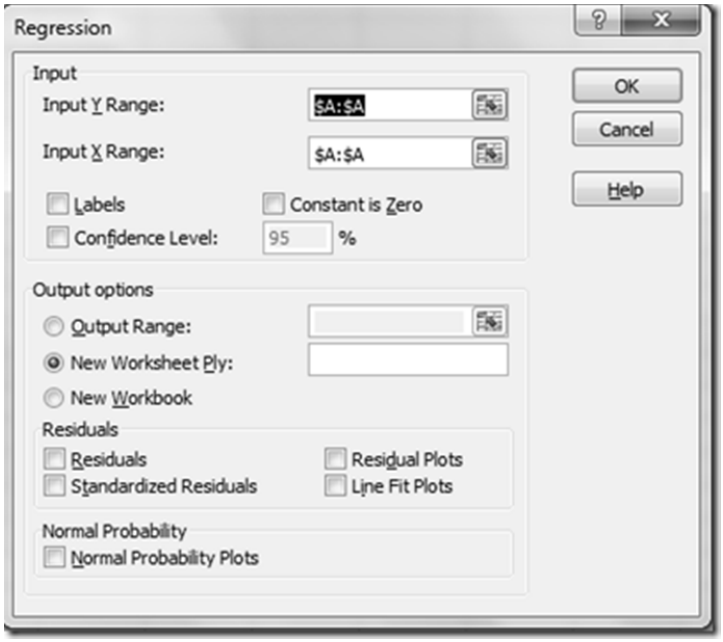
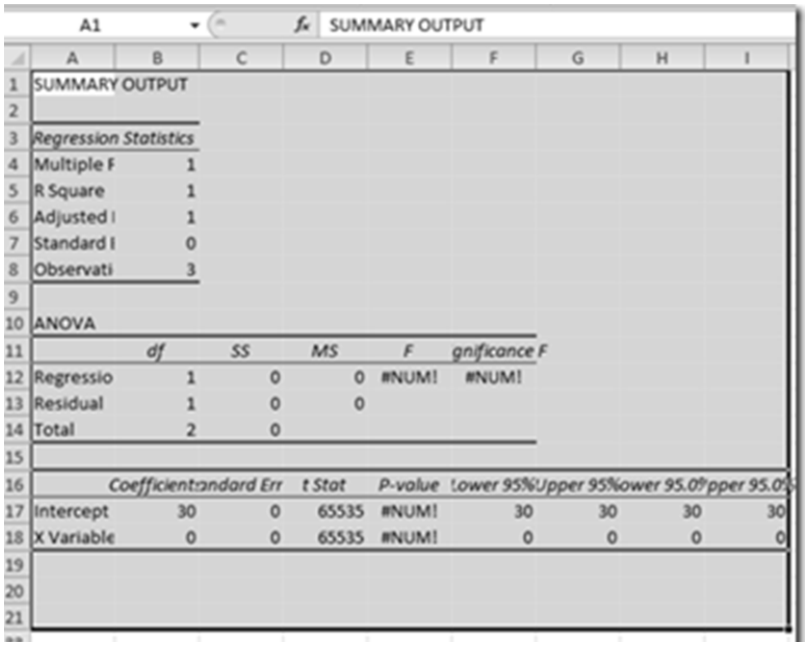


Fig. 3.65: Regression dialogue box



	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	*Regression Statistics*								
4	Multiple F	1							
5	R Square	1							
6	Adjusted I	1							
7	Standard I	0							
8	Observati	3							
9									
10	ANOVA								
11		*df*	*SS*	*MS*	*F*	*gnificance F*			
12	Regressio	1	0	0	#NUM!	#NUM!			
13	Residual	1	0	0					
14	Total	2	0						
15									
16		*Coefficient*	*tandard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
17	Intercept	30	0	65535	#NUM!	30	30	30	30
18	X Variable	0	0	65535	#NUM!	0	0	0	0
19									
20									
21									

Fig. 3.66: Output of regression

Macros

Macros are a feature of office which help you by automating easy repetitive tasks. It works by letting you record an action and then you can repeat the action as often as you like, with the click of a button. By default macros are disabled in Office 2010/2013, for enhancing security. Due to this, every time you open an office document containing macros (Word, Excel, etc.) office will give a warning and will require you to confirm to use Macros. This can be a hassle, however it is not very difficult to enable macros permanently. Click on the File menu, click Options, click on the Trust center tab. In this screen click on Trust center settings.

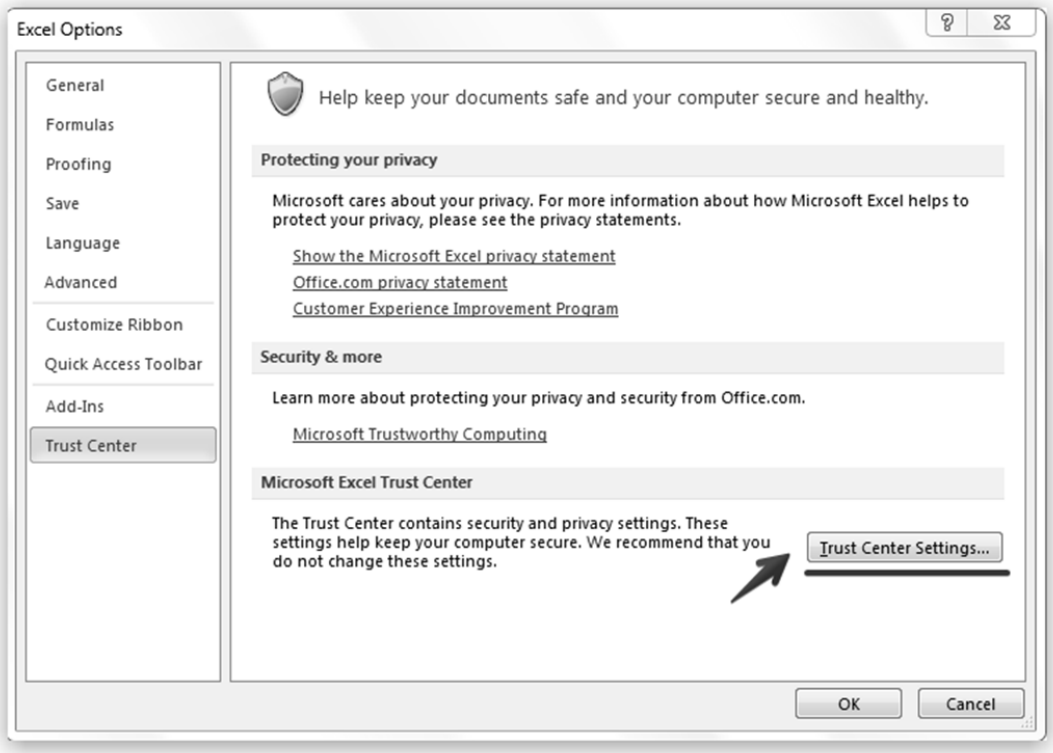


Fig 3.67 Trust Center Settings

In the Trust Center Settings screen first click on the Macros Settings tab. Then select the option: Enable all macros.

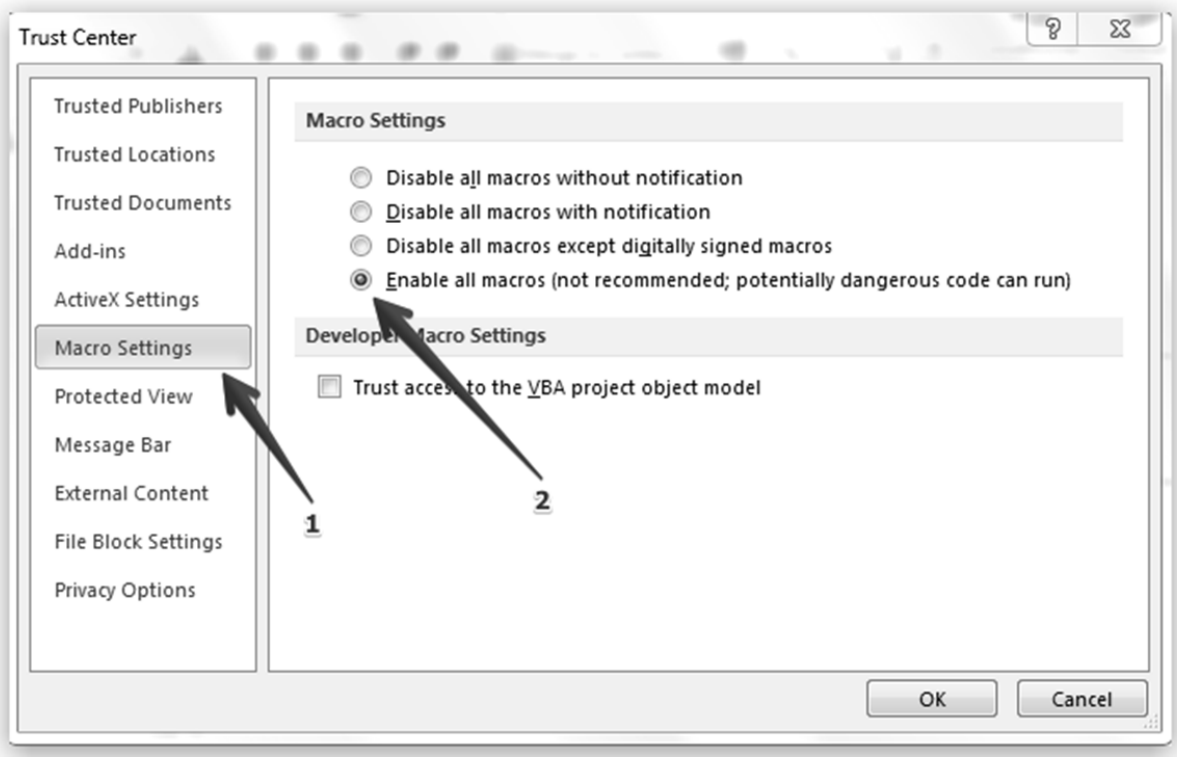


Fig. 3.68: Enabling Macros

Example: we have two worksheets containing marks of students and the grade they received. For one worksheet we will make a formula to calculate whether they have passed the course and we will use the macros function to apply it to the other sheet.

	A	B	C	D	E
1	ID	Name	Score	Final Grade	Status
2	1	Jan	50		
3	2	Jaap	60		
4	3	Kees	40		
5	4	Koos	70		
6	5	Pim	24		
7	6	Thijs	43		
8	7	Gerd	68		
9	8	Ruurd	72		
10	9	Joris	56		
11	10	Tim	34		

Fig. 3.69 A data sheet

Now we want to calculate the final grade based on the score on the test and then calculate whether the student has passed the course. We start with setting the macro function on Record. Click the View tab and then the macros dropdown button. Click Record Macro (see Figure 3.70).

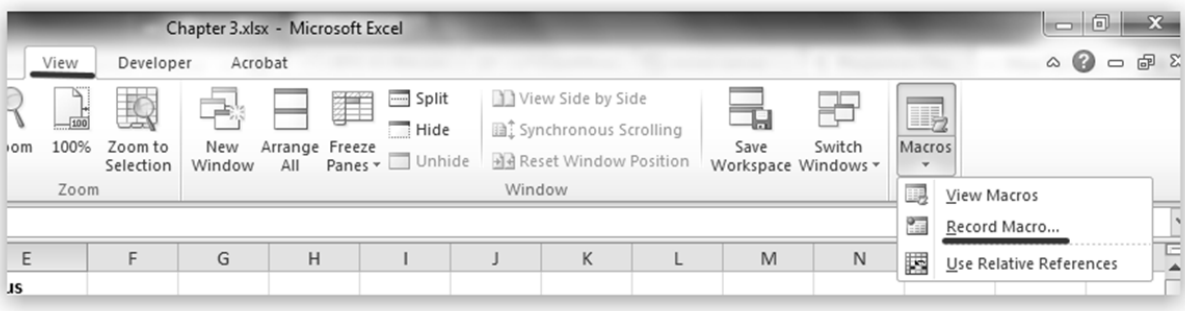


Fig. 3.70: Selection of Macro record option

Give the macro an appropriate name so you can remember what it does. You can also give it a shortcut key. This allows you to quickly run the macro again and again without having to click on the macros menu.

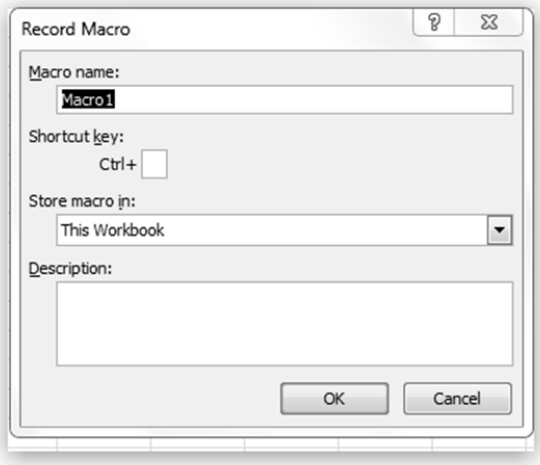
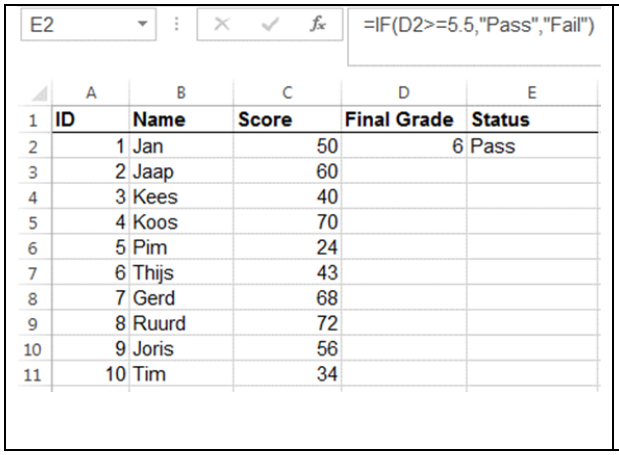


Fig. 3.71: Record a macro

Now that we have set the macro function to record we will enter our formula to calculate the Final Grade. The final grade in this example is calculated by Final grade = (exam score + 10)/10. So the MS-Excel formula for cell D2 becomes “=(C2+10)/10”.

The Status formula looks at whether the grade is a 5,5 of higher. In this case the student has passed the course, if the grade is lower the student has failed the course. Therefore the formula for cell E2 is “=IF(D2>=5,5;"Pass";"Fail")”.



	A	B	C	D	E
1	ID	Name	Score	Final Grade	Status
2	1	Jan	50	6	Pass
3	2	Jaap	60		
4	3	Kees	40		
5	4	Koos	70		
6	5	Pim	24		
7	6	Thijs	43		
8	7	Gerd	68		
9	8	Ruurd	72		
10	9	Joris	56		
11	10	Tim	34		

Fig. 3.72: Obtaining a student’s status with a macro

A	B	C	D	E
ID	Name	Score	Final Grade	Status
1	Jan	43	5,3	Fail
2	Jaap	68	7,8	Pass
3	Kees	72	8,2	Pass
4	Koos	56	6,6	Pass
5	Pim	34	4,4	Fail
6	Thijs	50	6	Pass
7	Gerd	45	5,5	Pass
8	Ruurd	67	7,7	Pass
9	Joris	79	8,9	Pass
10	Tim	43	5,3	Fail

Fig 3.73: Macro run on the second worksheet

Now that we have entered the correct formulas for the first row, we will drag down the formulas to apply them to all the rows below it (how to drag down a formula is explained in the section above). Now that we have applied the formula to all the rows for the final grade and for the status we will Stop Recording the macro, we are finished. Click stop recording in the Macros menu.

Now go to the other worksheet, which is exactly like this one, except for the fact that the final grade and status column are empty (you do not even have to fill in the formulas in the first row, as our macro has

recorded it all). Simply click on the macros menu, select your macro and press Run (or use the keyboard shortcut if you entered one). Automatically all the fields will be filled in!

3.8 Further spreadsheet study

There are excellent practical books written about spreadsheets which can help with the development of advanced use. We recommend Tennent and Friend's (Tennent & Friend, 2005) "Guide to business modeling" and Evans's "Business Analytics" (Evans, 2013). For a more advanced study we recommend books on decision support systems and business intelligence (e.g. (Turban, Aronson, & Liang, 2005)). There are several tools suppliers in the industry. For example: Microstrategy, www.microstrategy.com.

If you believe that your prior knowledge is insufficient, you can use online resources to learn more about Excel. Examples are:

- http://www.free-training-tutorial.com/
- http://www.excel-2010.com/

3.9 Exercises

1. Printing or e-readers? Imagine an editorial office with 250 people, who produce articles and read a lot for their research. At the moment people mostly read articles after having downloaded them and having them printed out. An average employee reads about 7 articles a week and reads also about 10 draft manuscripts from colleagues a week. An average article is 10 pages. There is only one very fast network printer for the whole group. This printer needs a new toner after 2,000 pages, costing 100 euros. One person is fully occupied managing this printer (salary costs 45000 per year). Printing out takes only 10 seconds, however, people have to walk to the printer (about 1 minute for each print out), and there is an average queue of 1 minutes per each print. Paper costs are about 10 euros per 500 pages. The average salary cost of the editorial office is about 60000 per year. Assume that one works 1600 hours per year. John Michels considered giving all employees an IPad for 600 euros a piece for free for each two years. He assumes that this will result in at least 75% less printing. Calculate what the costs and benefits are of this idea for the office. Also test this idea using different assumptions regarding % of print reduction, time spend on printing, and number of prints per person.

2. Marinus Jansen is a farmer with extensive experience in dairy production. The price he receives for raw milk has been declining in the last few years and consequently he considers of moving into the business of biological production of raw milk. The price he can receive for biologically produced raw milk varies between 20 and 40 euro cents per litre, depending on the demand on the

market and the size of supply by the producers (both are increasing at the moment). The price is also highly depending on the percentage of milk fat in each litre produced, because a higher fat percentage offers more opportunities for the manufacturer of creating all kinds of by products like cheese and butter. Each extra percentage fat above the average increases the price gained by one extra per cent. Marinus thinks that he will be able to produce milk with 20% more fat than the average, although depending on the weather conditions cows will have to be taken in stables resulting in much less fat. In bad situations even 20% less than the average can happen. The actual production costs are also highly depending in the weather conditions. Cold and rainy weather requires the cows to be on stable, and each day on the stable raises the costs of food and care per day with about 50 euros per cow. If additional feed is given to keep the level of fat per liter of milk at the high level of 20% above average, an extra 25 euros per cow has to be spend per day. Marinus plans to have 200 cows and has a fixed costs per stables, cleaning, maintenance and personnel of 150000 per year. An average cow deliver 25 litre per day, and for a full year about 8000 litre. Create the causal diagram that is applicable here and develop the simulations and scenarios needed for Marinus to decide if he wants to have an average salary for himself of 100,000 euro per year within a time span of 5 years.

3. Cash forecasting. SmileYou is a specialist fashion company, started by Mary-Ann, who was previously a top fashion model but quit the business a few years ago. Her designs are very popular and the business is growing fast. At the moment the return per year is 30 million, and the company grows each year over 20 percent in volume and returns. Production is outsourced to companies in China, Turkey and India, who all demand pre-payment before delivery. The relation between production costs and sales price is 0.5. The net sales costs (including logistics and warehousing) are about 30% of the sales price. The time gap between payment of producers and payment by customers is 3 months. Calculate the cash flow for needs for SmileYou for the next two years, given an 8% interest level and fixed costs of 10 million that increase each year by 10%. Also calculate alternative situations, e.g. higher and lower growth rates, and higher and lower interest rates. Also make graphics for the next 5 years regarding cash flow needs. How much of the growth will have to be funded by banks or stocks?
4. In a theme park the average number of guests per month of the high season is 100,000. The average number of guests per month during off-season is 70,000. The high season starts beginning of May and ends at August 31. The price of a ticket for the theme park is 22 Euros. There are two restaurants in the theme park. If you want to lunch in the park you can either pick a cheese or chicken sandwich, a pancake with sugar or chocolate or a pizza part. All the lunches are offered for 7 Euros each. If you want to dine in the park you can choose between 5 meals that

vary every month. A meal is offered at a price of 15 Euros. On average, 10% of the guests in the high season buy a lunch at the park and 5% of the guests dine at the park. The purchasing costs of the restaurant are 3 Euros for a lunch and 5 Euros for a meal. Furthermore, the park has a souvenir shops and candy stands (where you can also buy drinks). The average costs of purchasing the drinks and candy for the candy stands is 45,000 Euros per month. On average half of the average number of guests in a high season month buy something at the candy stand for 2 Euros.

The souvenir shops purchase for an average of 95,000 Euros per month. All the people who buy a lunch get a discount ticket for the souvenir shop. On average per month, all the people who buy a lunch buy something at the souvenir shop for 10 Euros. The souvenir shop sells for 100,000 Euros on average per month. The maintenance costs are 275,000 Euros on average per month and the theme park is investing for 300,000 Euros per month on average (which is seen as a cost in this context). The personnel costs are 500,000 Euros on average per month.

1. Make a causal diagram in Visio with the elements of the revenues and costs and how they influence the profit of the theme park.
2. Implement the causal diagram in MS Excel. Note that the revenues from the tickets are based on a month of high season.
3. The theme park wants to achieve a profit of 1,500,000 Euros next year. How many more guests this year does the theme park have to attract to break even.

5. An here is our theme question: How would you answer the question of how rationalism enables the collective brain? What knowledge or information is at stake and how can it become collective? What is the role of spreadsheets or other types of decision support systems in this context?

4. Analytic thinking and multi-perspective business modeling

4.1 Kantian inquiring system and multi-perspective modeling

Following 18th and beginning 19th century philosopher Kant, observing, describing and understanding anything (and thus also collecting information and data) is enabled by analytic a priori. All these analytic a priori's can be integrated into a more or less coherent analysis and design language, which we name synthetic a priori after Kant (Hartnack & Hartshorne, 1967). These synthetic a priori's often have a view of the world at its root, which we name an ontology (note for example that the word "profit" may comprise and exclude different categories depending on the fiscal policy of an organization and that competing firms will offer different activities for a similarly named service) and for business (process) modeling different software tools offer different tools and features that correspond with one essential understanding of the phenomon, i.e., its ontology (e.g. for business modeling ARIS, Bizagi and Petrinets offer very different sets of analytic a priori's). For analyzing organizational and business processes, several a priori's are important: time (when), space (where) (both dimensions are from Kant), what, how, who and why (these final four are from Sowa and Zachman (Sowa & Zachman, 1992)). Each model gives a different view of the phenomenon and the selection of the a priori's is therefore an important decision. Each decision here implies certain opportunities and limitations. Nowadays there are many software tools to describe and analyze a phenomenon by mapping and modeling different perspectives and then integrate these perspectives. Several of these perspectives are discussed in this chapter, and we offer a Kantian information management model as starting from scope definition, going to analysis using different a priori's, next aiming at integration of these perspectives, for the final purpose of supporting organizational change and business process improvement. See Figure 4.1.

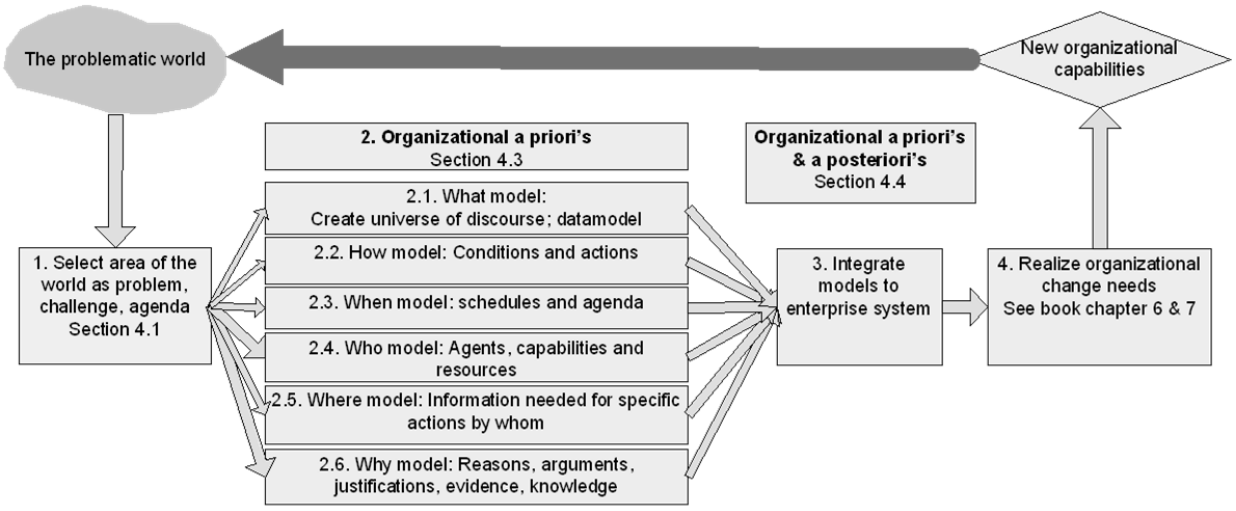


Figure 4.1: The Kantian information model (Wijnhoven, 2009b)

4.2 Goal of this chapter

People are able to make empirical observations, named "a posteriori propositions", following Kant, by using (analytic and synthetic) a priori as spectacles to observe the world. However, while these spectacles may enable viewing, they are necessarily biased and incomplete. These biases are not always a bad thing, because people are not able to observe everything and depending on scope and purpose some bias is required. But, if we are biased, we have to be aware that we are and thus intentionally should select a bias or understand it and adopt other perspectives to be integrated. This implies that the concepts we use for acquiring and organizing data (like data models and calculations models) are "a priori propositions" and that the data in our databases and the outcomes of our spreadsheet calculations are "a posteriori propositions".

In this chapter, we first describe elements of MS-Visio as a set of a priori's among which we will need capabilities to choose for being able to use them successfully in section 4.3. More specifically for organizational modeling (yes, we need models as representations, thus information, for storing and exchanging organizational insights!) we will use the following a priori's and related MS-Visio tools: "What models", "How models", "When models", "Who models", "Where models" and "Why models" with MS-Visio (the last being introduced in chapter 2). These are introduced in section 4.4. Next, we will integrate the models to an enterprise system, thus synthetic a priori understanding of organizations in section 4.5. After this introduction of a priori's, we discuss more complex integrated modeling languages and discuss one international standard language named Business Process Modeling Notation (BPMN). BPMN is not only a modeling language but there are tools available such as the BIZAGI BPMN modeler,

which also offer a software platform to automatically execute a process. By this, databases and business rules and logic are integrated into a new business process. We illustrate how this can be done in section 4.7.

4.3 More about MS-Visio and it's organizational a priori's

MS Visio is a business modeling tool by which one can create drawings, flowcharts and other reports of relevance for organizations and business. When you start with MS-Visio, the start screen of Figure 4.2 comes up. In this screen, you can choose which model you want to create.

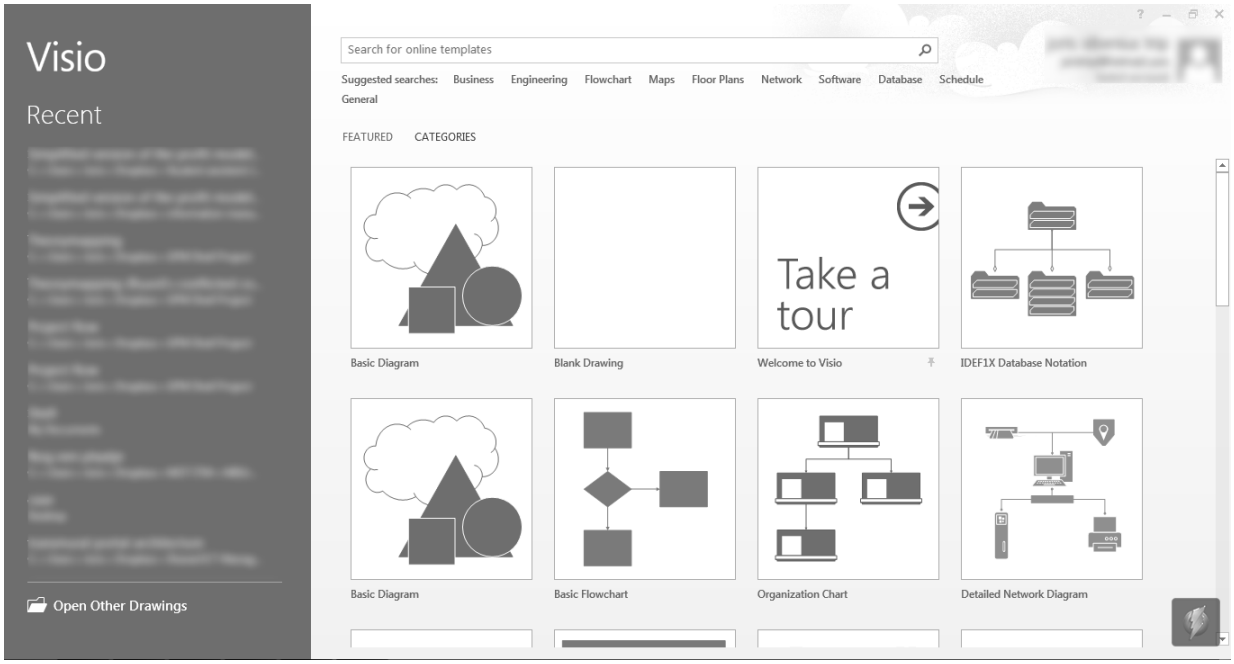


Figure 4.2: The MS Visio start screen

In addition to these standard templates, there is also the possibility to download templates from the Internet. You can do this by typing in the search bar at the top of the screen. It is also possible to choose another template by creating a new diagram. When you choose a template, it appears on the left side of the drawing page. This template (also named "stencil" by Visio) contains all shapes you need for a specific chart. Templates include, beside these shapes, also certain styles, settings and tools needed for that chart.

In this example we open a 'Floor Plan', which can be found under the category 'Maps and Floor Plans'. After opening it, we see our drawing environment with menus, toolbars and stencil shapes. The drawing

environment provides a 'grid', so it is easier to post shapes. Shapes contain connection points when you place them on the grid (drawing paper). By printing the drawing, the connection points and grid will disappear.

Suppose we want to make a building plan. The first step is to insert 'shapes'. If you open a template, you get an empty sheet, which is similar to the other Office programs. Now, draw by using different shapes. See Figure 4.3.

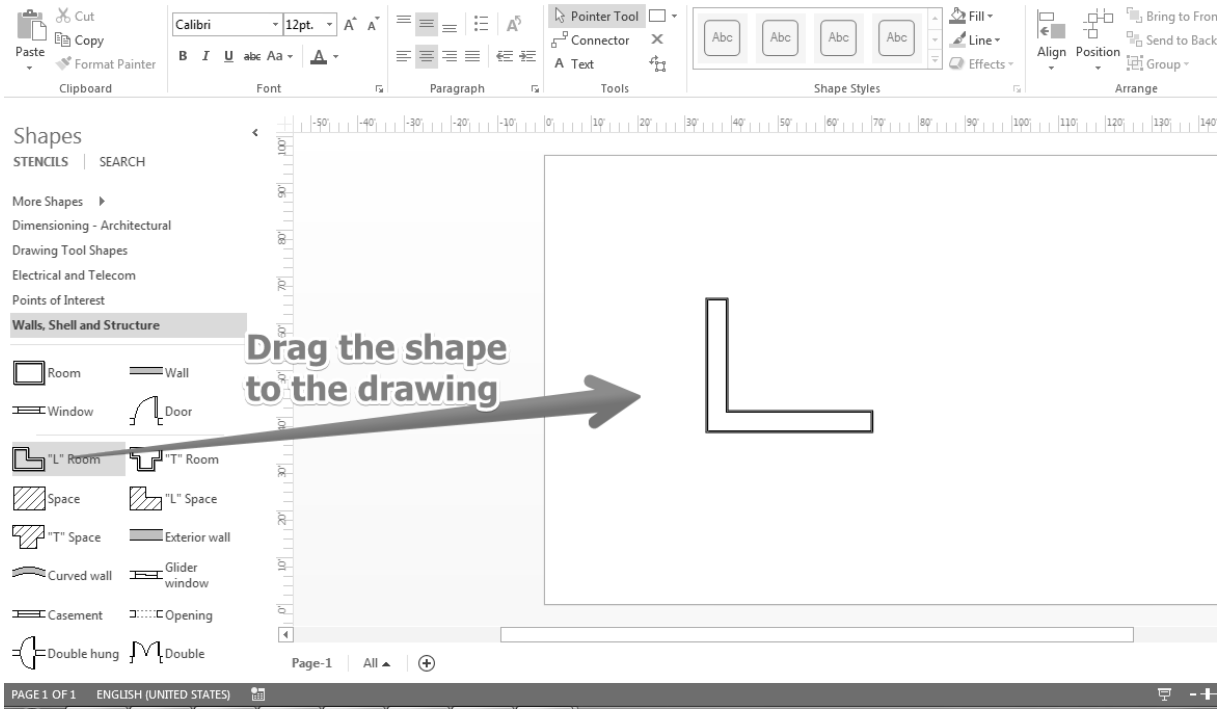


Figure 4.3. Selecting shapes

Sometimes you are looking for a shape which is not included in the list. Then you can use the search engine of MS-Visio, for finding that specific shape. Type in the search bar a description of the subject that you want to use and click on enter. The search bar is at above the shapes menu on the left.

One can also open a new stencil in the same template. Go in the menu bar to 'File', then 'Shapes', and find the shapes you want in the drawing page. If you added a shape in your drawing, you may also want to adjust it. There are two ways to do so:

1. Select the shape and move the mouse to one of the boxes, and pull the shape to the right size as given in Figure 4.4.

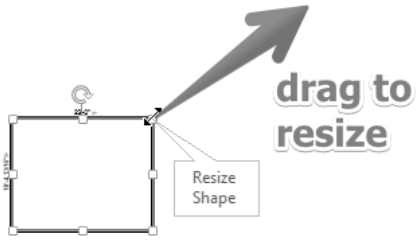


Figure 4.4: resize shape

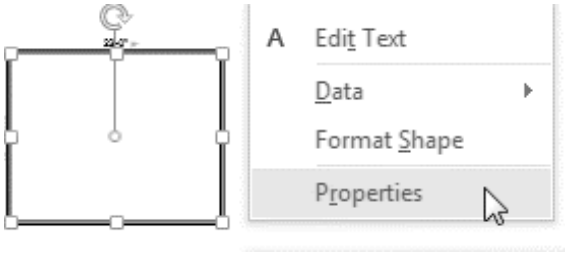


Figure 4.5: Change shape properties

2. Right click on one of the lines, then choose 'Properties' to specify the right format of the shape (see Figure 4.5).

Sometimes you would want different options: go to 'View', then choose 'Toolbars' and finally choose the toolbars you want at the top of your screen. Sometimes the template does not have the right size, which one can change by going to the bottom of the page see 'Page 1'. If you right-click on that you can see the option 'Page Setup'. In the different tabs you can specify exactly what you want. See Figure 4.6.

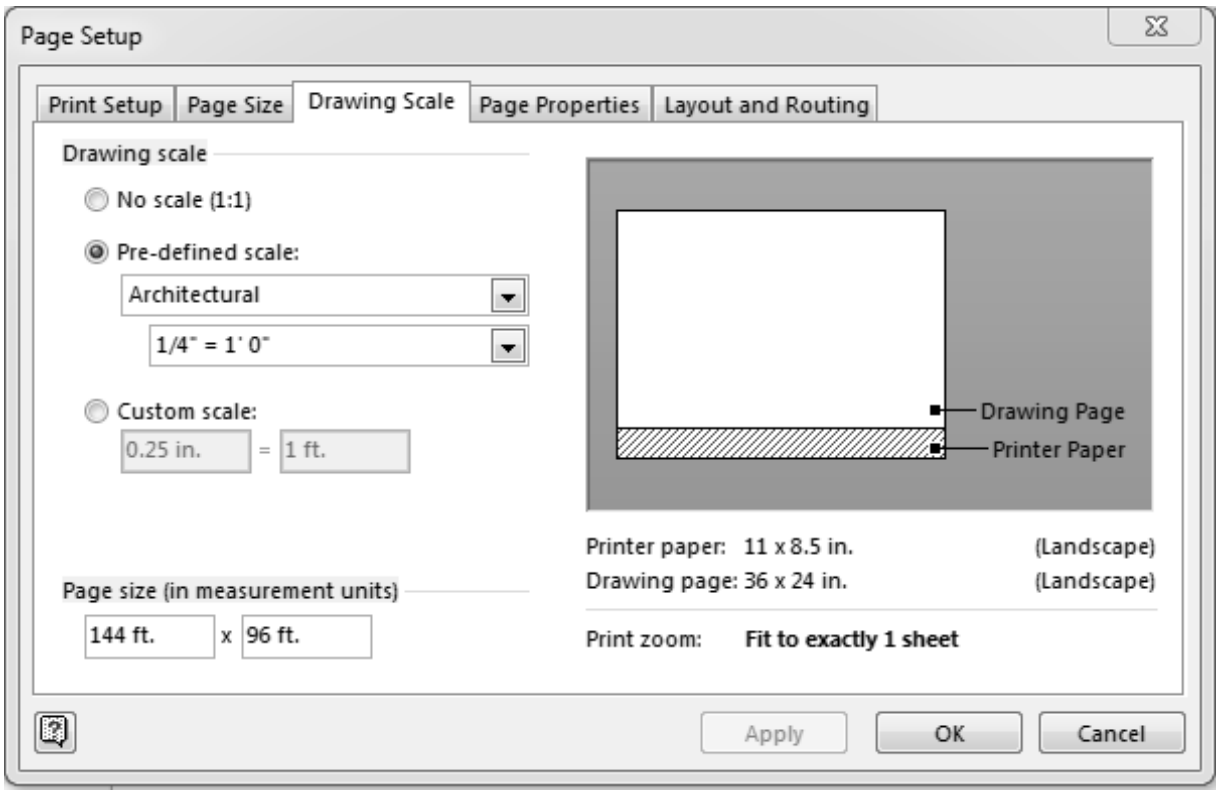


Figure 4.6 MS-Visio Page setup

As in any program, there are many Visio features that are invisible, but you can change this. Go to the 'Tools' button in the menu bar and then choose ‘Options'. See Figure 4.7.

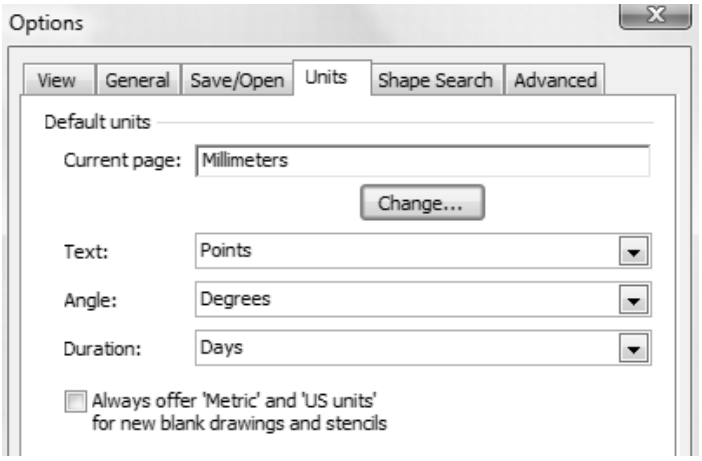


Figure 4.7: Options

Text can be added to the shapes. Click on a shape and start typing. MS-Visio zooms automatically to the shape, so you can read what you are typing. If you want to delete text quickly, double-click the shape. This makes the whole text in the shape selected. Then delete it, or type a new text. In addition to text in shapes you can insert text without a shape in MS-Visio. This can be done with a textbox. Just click on the text tool on the toolbar, and make a square or rectangle with a single click on the drawing page and then drag. Thereafter you can insert the text and edit it. For example you can underline or *italicize* the text.

For connecting different shapes and thus connecting a priori’s to a synthetized model, MS-Visio has the ‘Connector Tool’. Put your mouse on a junction of the shape, click and drag to the other shape to connect it. The junctions will be red, when the shapes are connected. Once they are connected, you can freely move the shapes around; the connection will remain.

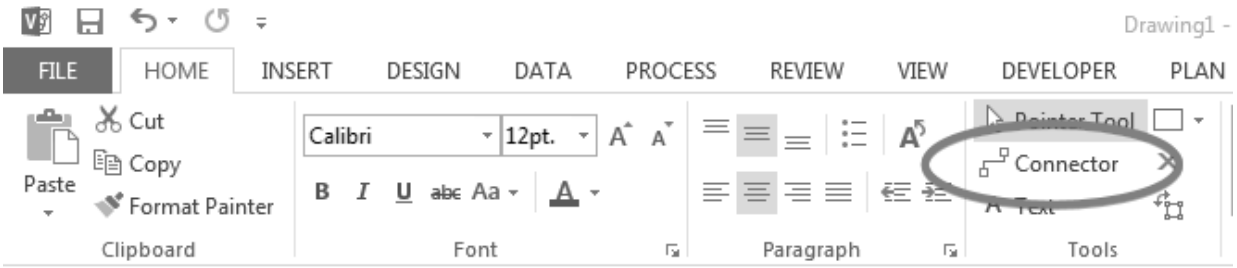


Figure 4.8: Connecting shapes

Frequently you will add the drawing to a MS-Word file. There are a few ways to insert a drawing in MS-Word.

1. You can save the drawing as a JPEG or Bitmap. Then open the Word file, click in the menu bar on 'Insert', and 'Picture'. Find the file on the location you saved it, and click on 'Insert'.

   When you save a shape as a JPEG, you cannot change that shape in MS-Visio anymore. Per default, MS-Visio saves the shapes as a drawing (.vsd). So when you want to change the shape in the future, you must save your file with the .vsd extension.
2. The other way is to select the model, or a part of it, and use 'copy' and 'paste' to insert it in the MS-Word file. The advantage of this is that you can change the drawing in MS-Word by a 'double-click'.

4.4 Business process modeling using MS Visio.

As marked in the introduction of this chapter, Kant's epistemology emphasizes different a priori's which result in different models. One can create the models in separate MS-Visio pages of the same 'Visio book'. For the different perspectives we need different templates. For instance a certain shape like a process, may belong to a template such as a flowchart. Following Kantian reasoning, a business process would be a "synthetic a priori", consisting of multiple "analytic a priori's" like activity, time, sequence, information flow and actor (depending on the ontology, of course more can be included or excluded). The template represents a 'synthetic a priori', in which the various shapes are related. A business process thus may be modeled from different perspectives (How, When, etc but also different actors limited views). Each of these perspectives has a ontology of the business process. Let us further describe some of these synthetic a priori's and how MS-Visio can be useful here.

'How' Perspective

The How perspective is a roadmap that describes what must be done in certain situations. For this perspective we choose the Basic Flowchart template under MS Visio "Flow Chart". Name the page before you start. You can re-name a page in the same way as an MS-Excel-sheet. In the bottom of your screen you see the tab 'Page', click with the right mouse-button, and type a proper name (e.g. how) at 'rename page'. The next step is to create a model by connecting the right shapes. A brief legend of the shapes is needed to describe the meanings of the shapes, like in Figure 4.9.

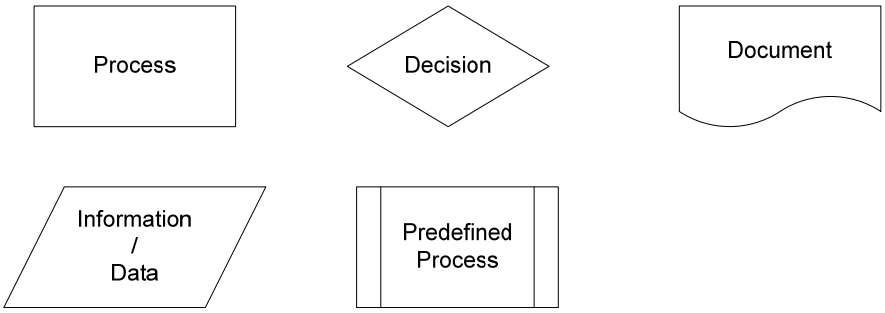


Figure 4.9. Key shapes (representations of a priori's) of the How perspective

An example of a "how model" is given in Figure 4.10.

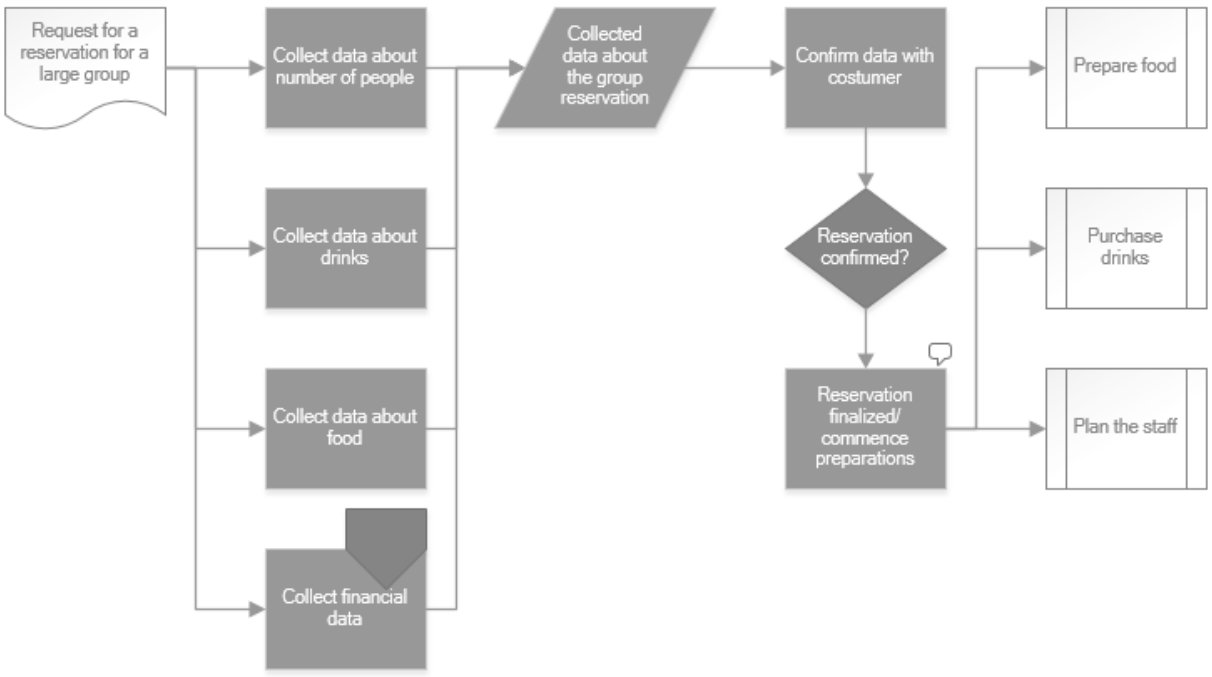


Figure 4.10 Example of a 'how' model for a reservation process

In the model are 2 unusual objects namely (1) the 'off page reference', which when double-clicked links directly to another model (This is discussed this later with the integrated workflow model) and (2) the commentary or a remark object, which can be created by clicking on 'Insert' in the menu bar, then

clicking on ‘Comment’. For both, see Figure 4.11 and 4.12. One can also add colors to the shapes to clarify the meaning of the different shapes.

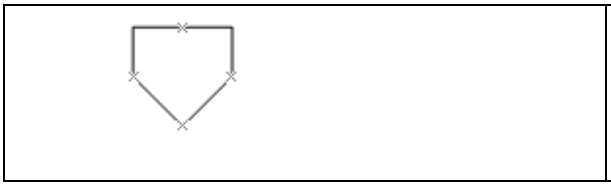


Figure 4.11 Off page reference

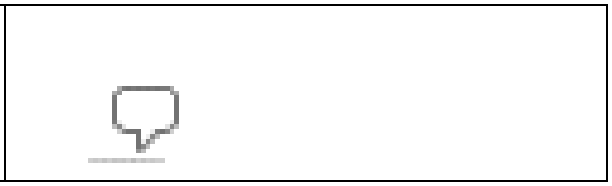
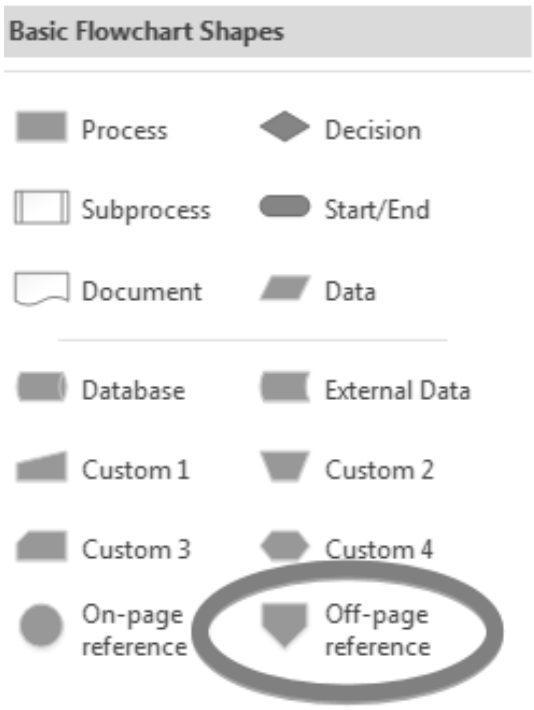


Figure 4.12: Comment

The off-page reference shape can be used to link full pages, or to link certain shapes from one perspective to a similar shape in a different perspective. In the first case it does not matter where the off-page reference shape is placed. In the second case, the off-page reference must be placed on or beside the appropriate shape. How do we get these off-page references inserted in the workbook? The ‘Off-page reference’ is situated under the ‘Basic Flowchart’ template. See Figure 4.13.



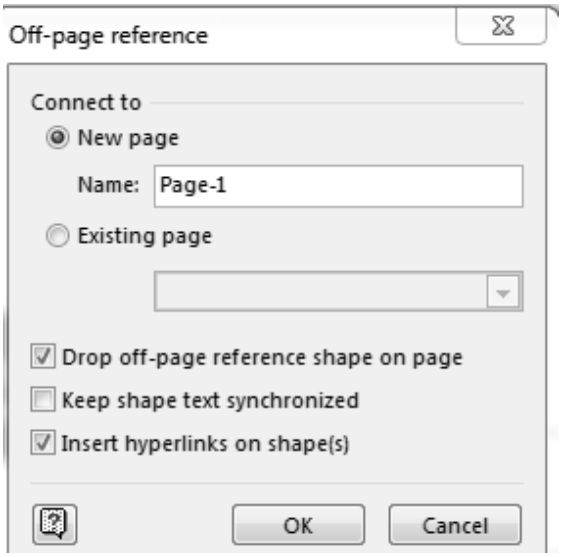


Figure 4.13: Selection of Off-page reference in a “basic flow chart”

Figure 4.14 Off page reference dialogue box

Click on it, and drag it to the right place and the next screen appears (see Figure 4.27). To link the different perspectives to each other, choose for ‘Existing page’ under ‘Connect to’. Then indicate to which page the link should be made. Do not change the other marks, and click on ´OK´. On the page, chosen under ´Connect to´, will appear the same ‘Off-page reference’. When you connect two shapes, drag the icon to the appropriate shape. Now the connection is done. What have we achieved? Double click on the off-page reference shape. Right now, you see the shape/page to which the off-page reference is attached.

‘Where’ Perspective’

The 'where' perspective shows the communication needs between departments if a project is under construction. The 'where' perspective consists of two parts. So open two new pages in your ms-Visio-workbook (’Insert’ → ‘New Page’ → ‘Name’).

The first step of the 'where' model is to create an ‘Organization Chart’ in which the sections are displayed briefly. Just make sure you choose the right template namely ‘The Organization Chart Template’. For opening this template click on ‘File’ in the menu bar, followed by ‘Shapes’, then ‘Business’ and click on ‘Organization Chart’. See Figure 4.15 for an example of an Organization Chart.

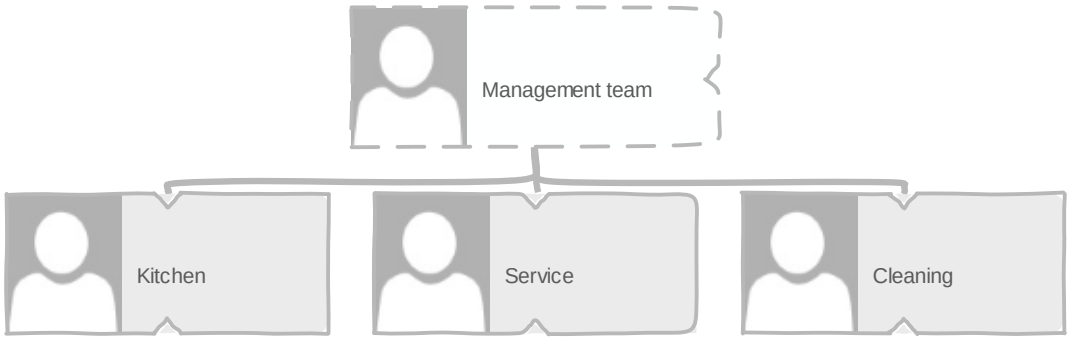


Figure 4.15: Organization chart

The second part of this perspective consists of an integrated model, which related the first part of the ‘where’ perspective to the ‘how’ perspective. Create it by using “swimlanes”, which represent the organizational unit responsible for one or more activities of the business process. Open a new page in the corresponding set of shapes by clicking on 'File' → ’Shapes’ →’Flow Chart’ → ‘Cross functional flowchart shapes’. Drag the ‘Swimlane’ on the drawing page.

You can either use vertical or horizontal swimlanes. You define the number of departments in the first part of this perspective. Use a swimlane for every department and name them. It is now just a question of exactly replicating the ‘how’ model, and decide which department makes which step. So copy the ‘how’ model, and place each step in the right department. Connect the shapes in the same way as in the ‘how’ model and see the result in Figure 4.16.

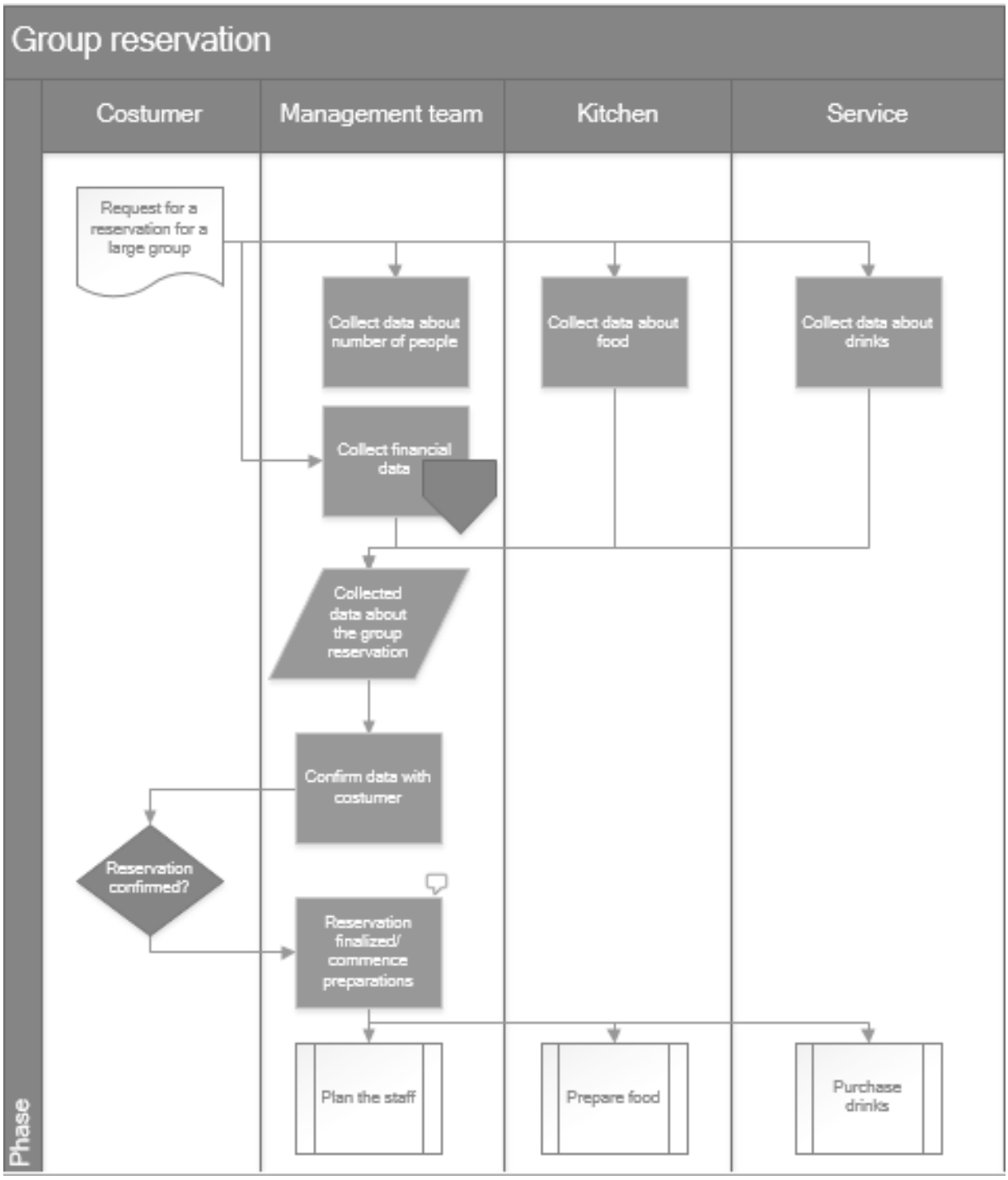


Figure 4.16: Example of a Where model (cross functional flow diagram)

‘Who’ Perspective

The 'who' perspective shows the interrelationships between actors, activities and resources. For modeling this perspective, we use MS Visio's 'Basic shapes' template ('General' → 'Basic Shapes'). The meaning of the shapes are given in Figure 4.17.

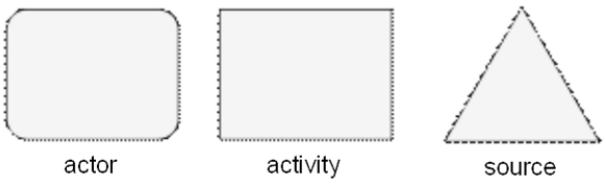


Figure 4.17: Key shapes of the Who model

The idea is as follows: an actor performs an activity, for which s/he needs some resources. Choose the 'Basic Shapes' template to represent this. See Figure 4.18.

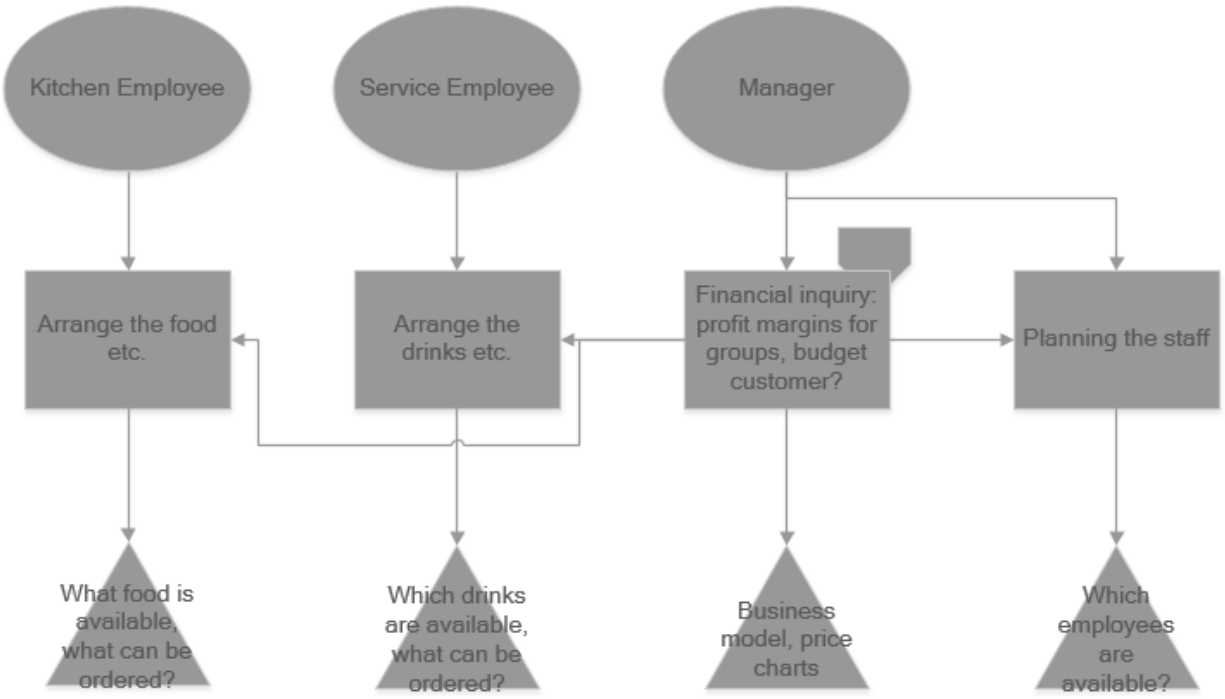


Figure 4.18. Example of a Who model

'What' Perspective

The 'what' perspective considers the needed data and information. For this perspective we need an Entity-Relationship Diagram, which we have seen in chapter 2. An alternative, non-MS-Visio, way of representing knowledge objects provides CMAP, see at cmap.imhc.us and Figure 4.19.

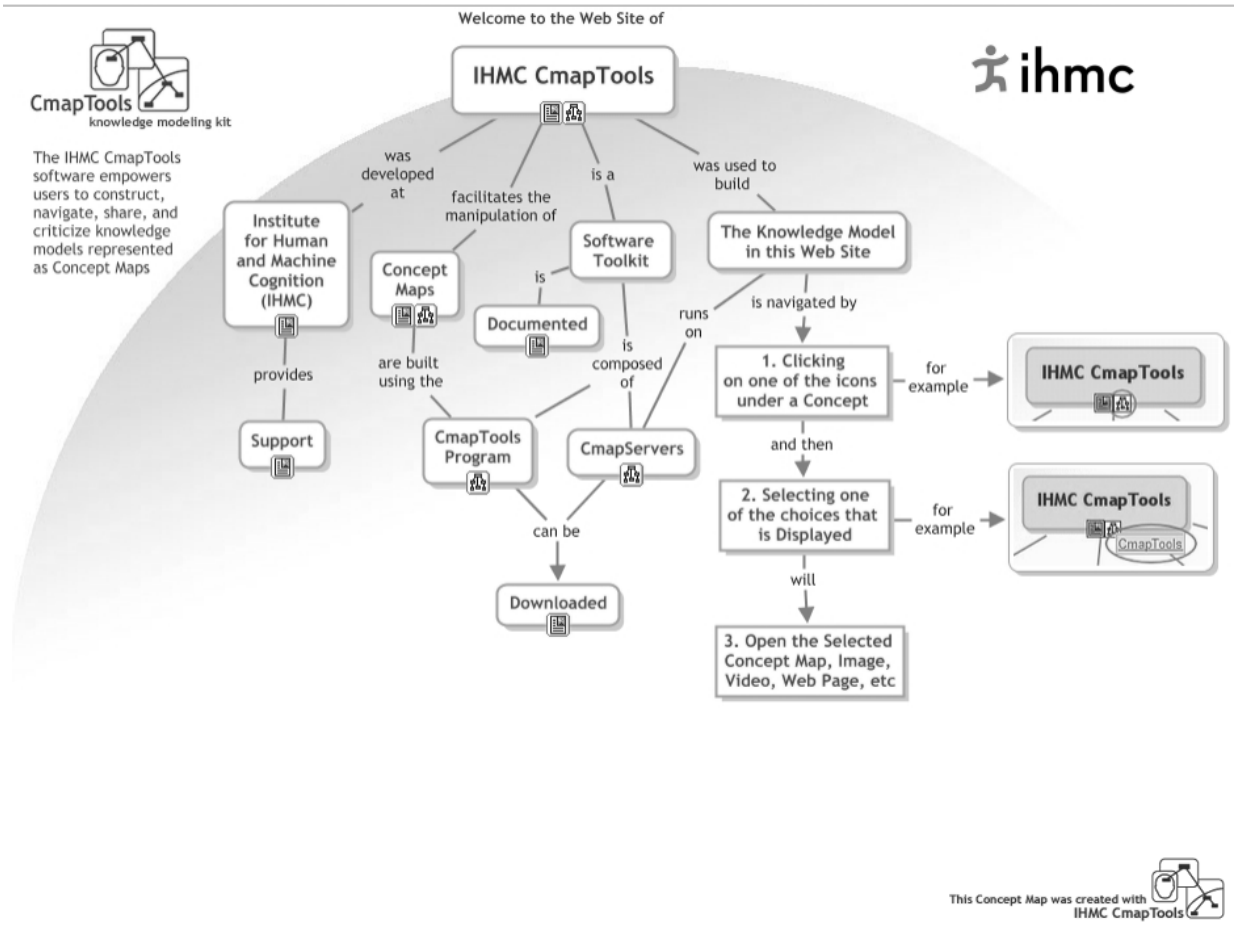


Figure 4.19.: CMAP

CMAP presents itself as follows: “The IHMC CmapTools program empowers users to construct, navigate, share and criticize knowledge models represented as concept maps. It allows users to, among many other features, construct their Cmaps in their personal computer, share them on servers (CmapServers) anywhere on the Internet, link their Cmaps to other Cmaps on servers, automatically create web pages of their concept maps on servers, edit their maps synchronously (at the same time) with other users on the Internet, and search the web for information relevant to a concept map.

CmapTools is used worldwide in all domains of knowledge and by users of all ages to graphically express their understanding. In particular, CmapTools is used in schools, universities, government organizations, corporations, small companies, and other organizations, both individually and in groups, for education, training, knowledge management, brainstorming, organizing information, among other applications. The collaboration and publishing features provide a powerful means for representing and sharing knowledge.

The IHMC CmapTools client is FREE for use by anybody, whether its use is commercial or non-commercial. In particular, schools and universities are encouraged to download it and install it in as many computers as desired, and students and teachers may make copies of it and install it at home."

Although CMAP is an interesting and useful tool for organizing knowledge (therefore CMAP models are sometimes also named "ontologies", which is not the same as the nature definition of a concept as we use it in this text) and brainstorming, it clearly has no links to a business process and thus it does not have a conceptualization of a business process as its underlying ontology.

'When' Perspective

As the name suggests, the 'when' perspective is about modeling of time dimensions, which can be done by PERT and Gantt charts in MS-Visio. Add for this perspective also a new page in your current workbook, and name it. Then open the correct template. For the 'Gantt'-chart, click on 'File' → 'Shapes' → 'Schedule' → 'Gantt Chart Shapes'. The following screen will appear, see Figure 4.20.

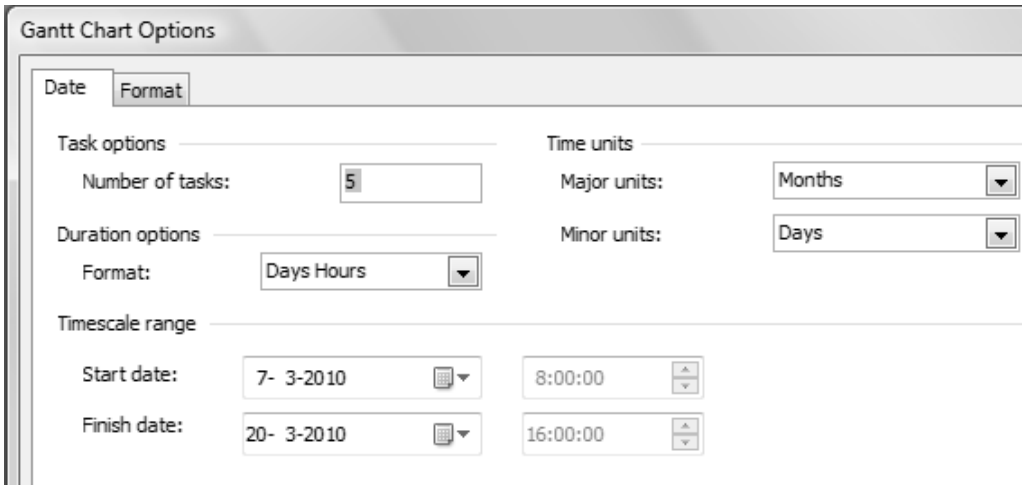


Figure 4.20: Gantt chart dialogue box

Before all kinds of options are introduced, we must think about why the 'when' perspective should be made. Is this an example of a project with a number of tasks, or is this a whole year with all sorts of projects? So think about the purpose. It is even better when a Gantt chart is made for both options.

When we choose to display the timing of a project, we must first determine how many jobs this project encompasses. Suppose there is a project with six tasks, we fill in “6” by 'Number of tasks'. Then we set the time dimensions. For a project the major units may be months, and minor units are days or possibly weeks. We must also determine in which time units the duration of a task is kept. Finally, we indicate when the project starts and ends. When everything is filled out click on 'OK' and we have an ‘empty’ Gantt chart as given in Figure 4.21.

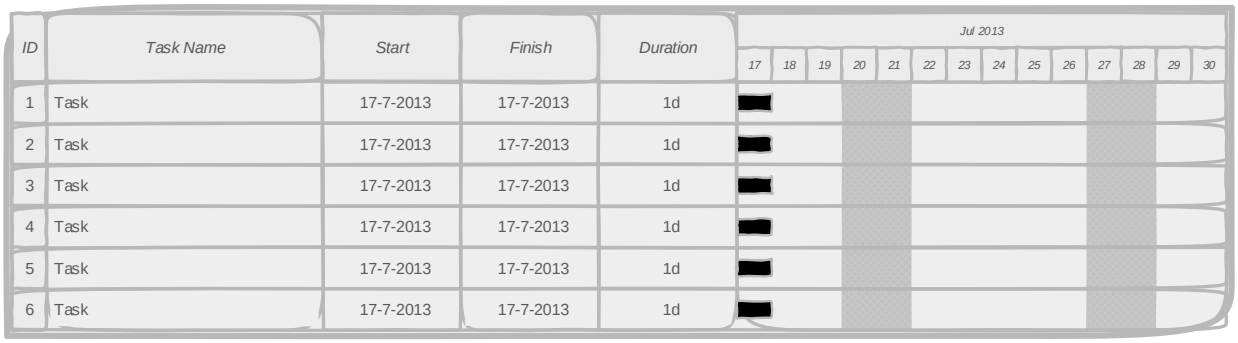


Figure 4.21: Empty When model of Gantt chart

Now we must fill in the correct information for a task. Double-click on ‘Task 1’ and enter the correct job name. Then double-click on the start date of this task, and change it in the correct date. Do the same for the end date. The ‘Duration’ bar behind the task will change automatically. Do this for all tasks. Tasks may have possible overlaps, however, some tasks can only be done when the previous task is finished. So you must think about the tasks, which of them are sequential? An example of a Gantt Chart is in Figure 4.22.

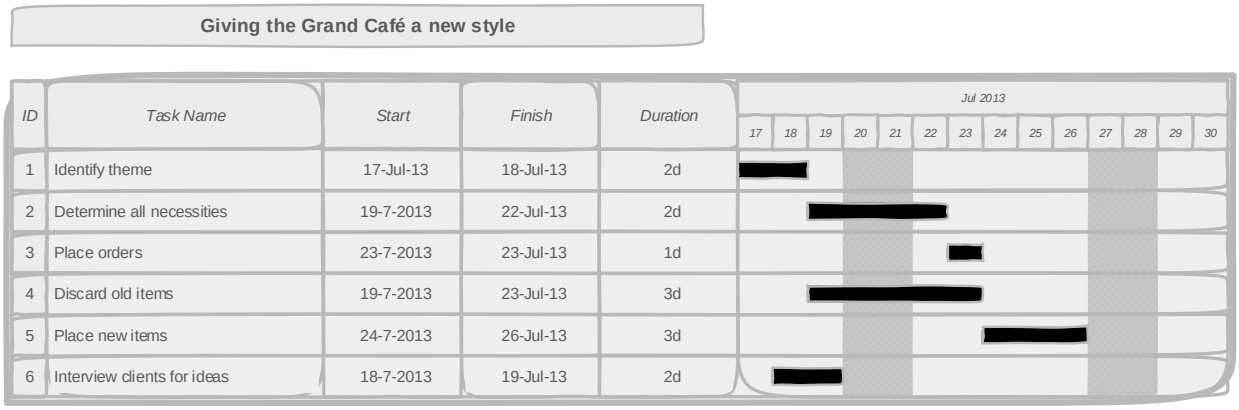


Figure 4.22. Example of a Gantt chart

Also a Title bar is added to this chart, which immediately gives an explanation of the purpose of the table. The 'Title bar' can be found in the 'Shape bar' on the left side of the screen. It is also possible to add individual tasks, when there are not enough tasks in the overview. Click on the 'Task bar', and drag it to the chart. The further steps are the same as above.

There are several other easy to use webservices that help in developing project plans and appointments, like tom's planner www.tomsplanner.com and www.doodle.com respectively.

'Why' Perspective

The last perspective is the 'why' perspective. This perspective seeks to provide evidence, arguments or reasons for certain actions. One can use the 'Brainstorming shapes' of Visio. Add a new page to your workbook, and name it. Then open the right template, the brainstorming shape, click on 'File' → 'Shapes' → 'Business' → 'Brainstorming' → 'Brainstorming shapes'.

This is a great way to make an informed decision, which weighs the pros and cons against each other. Add a 'central topic' to the page for the arguments pro, and add a 'central topic' for the arguments against a particular topic. Then add these arguments to both 'central topics'. Try to find at least 3 pros en 3 cons. Connect these arguments with arrows to the 'central topic'. An example of this perspective is given in Figure 4.23.

Figure 4.23: Why (Brainstorm) model

A very nice web service that enables group brainstorming is www.spilter.com.

We may conclude at the end of this section that many ways of modeling business reality and thinking exist. To avoid that this models become a disconnected chaos, we can connect and integrate them. This can be done in an ad hoc way by MS-Visio's off page reference, but several business and organization modeling tools use a modeling language which includes a specific and coherent understanding of objects in reality and the mind, named ontologies. Because these ontologies have an explicit description of categories and relations among them, modeling and organizational part by such a language and tool enables them to be executed by software. The next section gives such an explicit business modeling language (BPMN) and a software tool (Bizagi) which allows the execution of the resulting models by web services.

4.5 Ontologies, the BPMN language

The Business Process Management Initiative has developed a standard Business Process Modeling Notation (BPMN). A notation is the language by which the key constructs of an ontology can be represented. The primary goal of the BPMN effort was to provide a notation that is readily understandable by all business users, from the business analysts that create the initial drafts of the processes, to the technical developers responsible for implementing the technology that will perform those processes, and finally, to the business people who will manage and monitor those processes. BPMN is also supported

with an internal model that enables the generation of computer executable business processes. Thus, BPMN creates a standardized bridge for the gap between the business process design and process implementation. For the explanation of BPMN, we have used the Introduction to BPMN by Stephen A. White, published on the website of the Object Management Group Business Process Model and Notation (http://www.bpmn.org/). If you want to read more on the subject we recommend that you visit this BPMN website.

BPMN defines a Business Process Diagram (BPD), which is based on a flowcharting technique for creating graphical models of business process operations. A Business Process Model, then, is a network of graphical objects, which are activities (i.e., work) and flow controls that define their order of performance.

The BPD elements enable the development of diagrams that will look familiar to most business analysts. The elements were chosen to be distinguishable from each other and to utilize shapes that are familiar to most modelers. For example, activities are represented by rectangles and decisions are represented by diamonds. It should be emphasized that one of the drivers for the development of BPMN is to create a mechanism for creating business process models, while at the same time being able to handle the complexity inherent to business processes. The approach taken to handle these two conflicting requirements was to organize the graphical aspects of the notation into specific categories. This provides a small set of notation categories so that the reader of a BPD can recognize the basic types of elements and understand the diagram. Within the basic categories of elements, additional variation and information can be added to support the requirements for complexity without dramatically changing the basic look-and-feel of the diagram. The four basic categories of elements (synthetic a priori's) are named Flow Objects, Connecting Objects, Swimlanes, and Artifacts.

The three Flow Objects are (see Figure 4.24 and 4.25):

- An Event is represented by a circle and is something that "happens" during the course of a business process. These Events affect the flow of the process and usually have a cause (trigger) or an impact (result). Events are circles. There are three types of Events, based on when they affect the flow: Start, Intermediate, and End.
- An Activity is represented by a rounded-corner rectangle and is a generic term for work. An Activity can be atomic or nonatomic (compound). The types of Activities are: Task and Sub-Process. The Sub-Process is distinguished by a small plus sign in the bottom center of the shape.
- A Gateway is represented by the diamond shape and determines traditional decisions, as well as the forking, merging, and joining of activity sequence paths.

The Flow Objects are connected together in a diagram to create the basic skeletal structure of a business process (See Figure 4.24). There are three Connecting Objects that provide this function. These connectors are:

- A Sequence Flow, which is represented by a solid line with a solid arrowhead. Sequence Flow presents the order that activities will be performed in a Process. Note that the term "control flow" is generally not used in BPMN, but common in another process modeling language named event driven process chains (Scheer, 2000).
- A Message Flow, which is represented by a dashed line with an open arrowhead and is used to show the flow of messages between two separate Process Participants (business entities or business roles) that send and receive them. In BPMN, two separate Pools in the Diagram will represent the two Participants.
- An Association, which is represented by a dotted line with a line arrowhead and is used to associate data, text, and other Artifacts with flow objects. Associations are used to show the inputs and outputs of activities.

For modelers who require or desire a low level of precision to create process models for documentation and communication purposes, the mentioned core elements plus the connectors will provide the ability to create understandable diagrams like Figure 4.24.

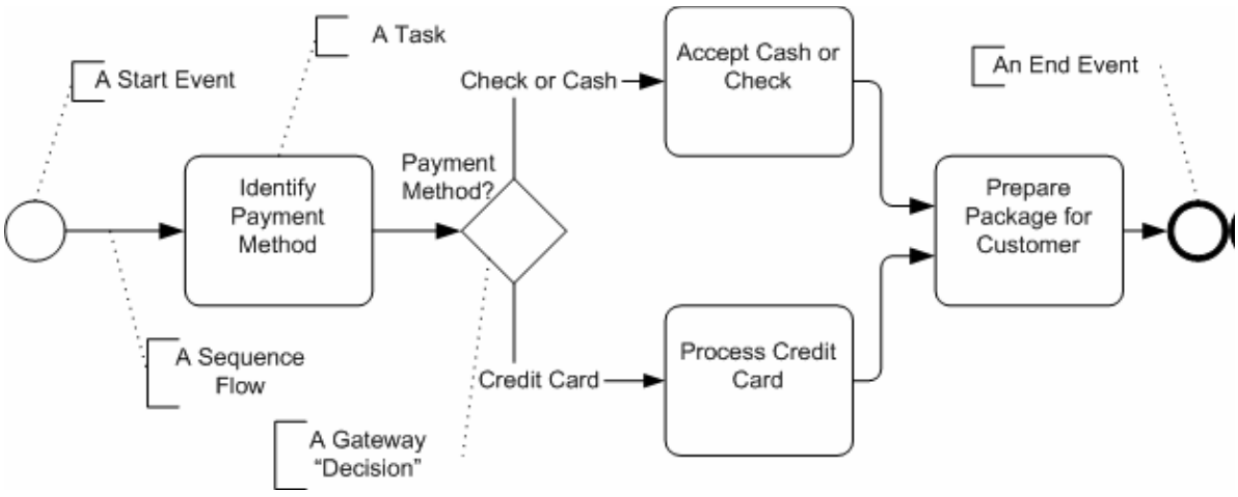


Figure 4.24: An Example of a Simple Business Process

For modelers who require a higher level of precision to create process models, which will be subject to detailed analysis or will be managed by Business Process Management System (BPMS), additional details can be added to the core elements by internal markers (see Figure 4.25).

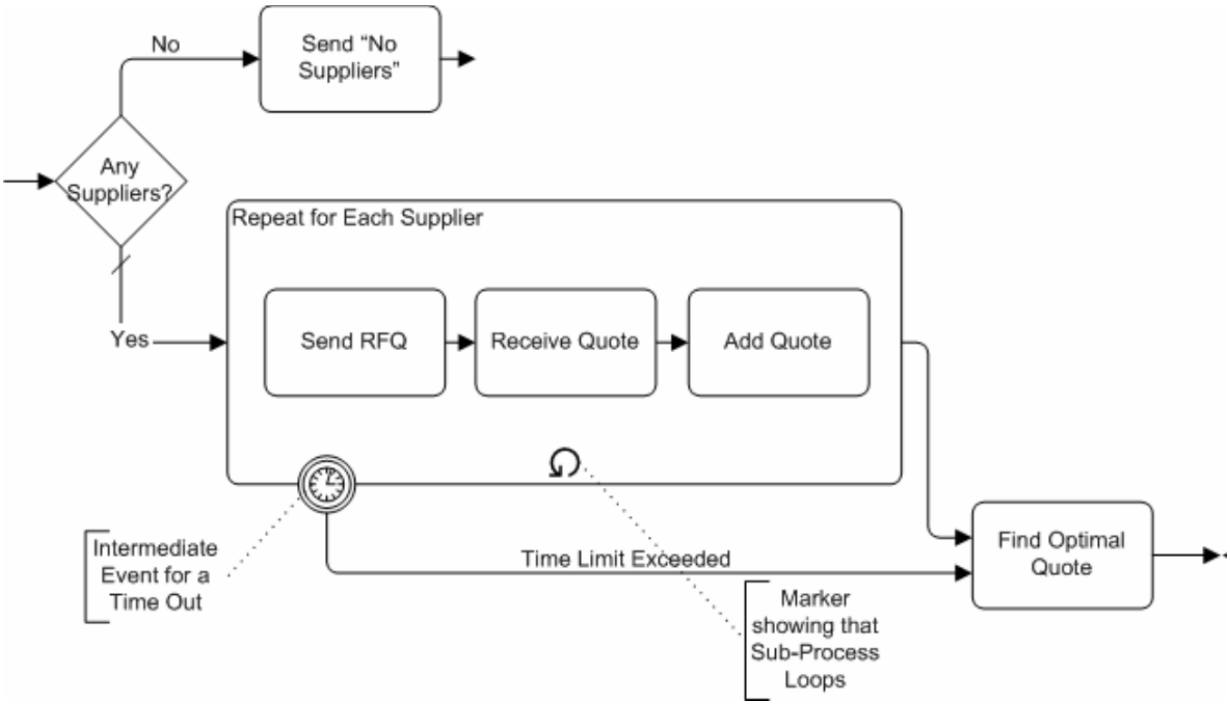


Figure 4.25: A Segment of a Process with more Details

Many process modeling methodologies use the concept of swimlanes (also named cross functional flow diagrams). BPMN supports swimlanes with two constructs…

- A Pool, which represents a Participant in a Process. It also acts as a graphical container for partitioning a set of activities from other Pools, usually in the context of B2B situations (see Figure 4.26).
- A Lane, which is a sub-partition within a Pool and will extend the entire length of the Pool, either vertically or horizontally (see Figure 4.26). Lanes are used to organize and categorize activities.

Pools are used when the diagram involves two separate business entities or participants (see Figure 4.26) and are physically separated in the diagram. The activities within separate Pools are considered self-contained Processes. Thus, the Sequence Flow may not cross the boundary of a Pool. A Message Flow shows the communication between two participants, and, thus, must connect between two Pools (or the objects within the Pools).

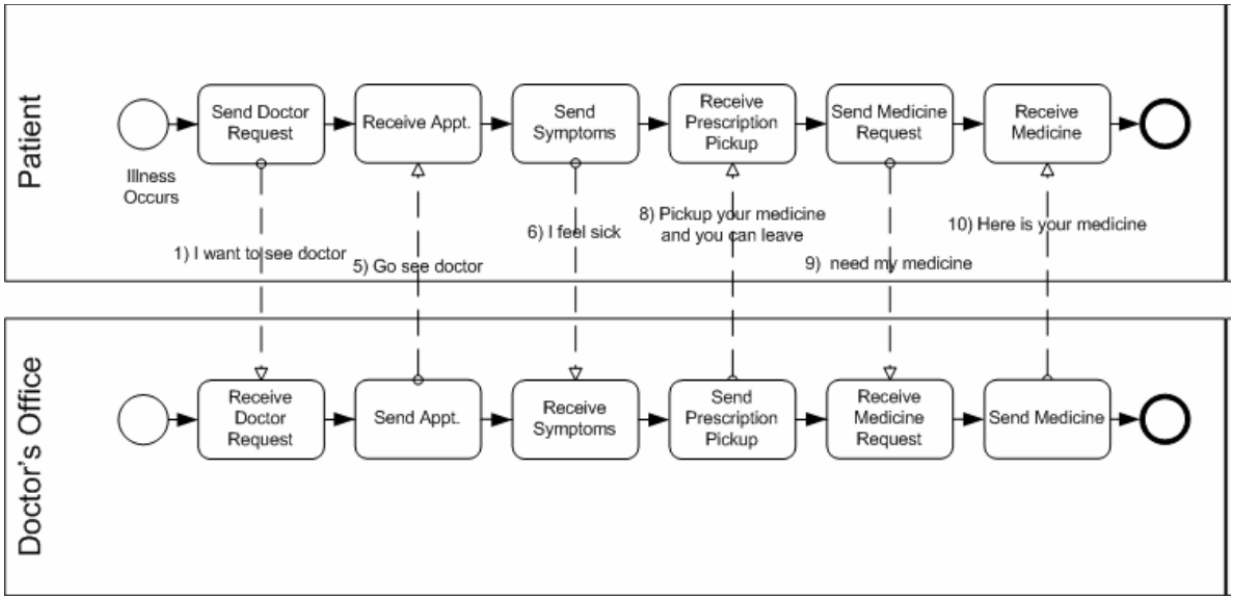


Figure 4.26: An Example of a BPD with Pools

Lanes are more closely related to the traditional swimlane process modeling methodologies. Lanes are often used to separate the activities associated with a specific company function or role (see Figure 4.27). Sequence Flow may cross the boundaries of Lanes within a Pool, but Message Flow may not be used between Flow Objects in Lanes of the same Pool.

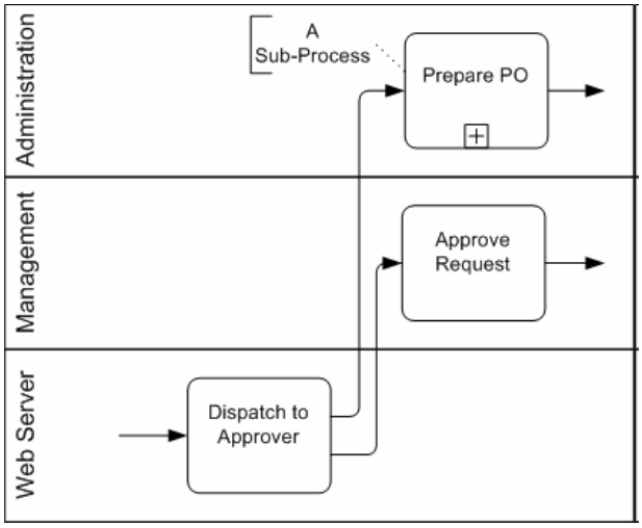


Figure 4.27: A Segment of a Process with Lanes

Any number of Artifacts can be added to a BPMN diagram as appropriate for the context of the business processes being modeled. The current version of the BPMN specification pre-defines only three types of BPD Artifacts (see Figure 4.28), which are:

- Data Objects are a mechanism to show how Data Object data is required or produced by activities. They are connected to activities through Associations.
- A Group is represented by a rounded corner Group rectangle drawn with a dashed line. The grouping can be used for documentation or analysis purposes, but does not affect the Sequence Flow.
- Annotations are a mechanism for a modeler to provide additional text information for the reader of a BPMN Diagram and are represented in BPMN by a text preceded by [Sign.

Modelers can create their own types of Artifacts, which add more details about how the process is performed—quite often to show the inputs and outputs of activities in the Process. However, the basic structure of the process, as determined by the Activities, Gateways, and Sequence Flow, are not changed with the addition of Artifacts in the diagram; as you can see by comparing Figure 4.27 and Figure 4.28.

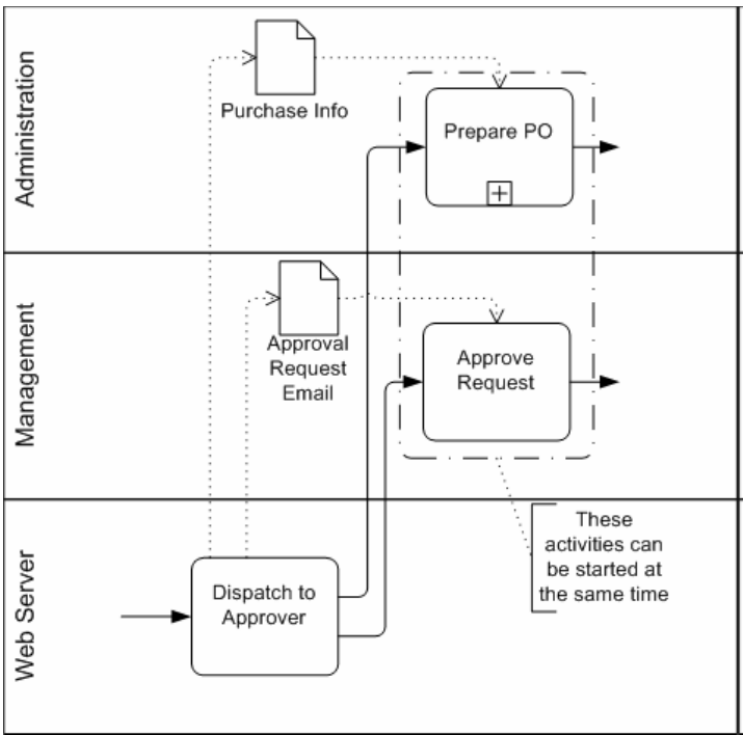


Figure 4.28: A Segment of a Process with Data Objects, Groups, and Annotations

Within the variety of process modeling objectives, there are two basic types of models that can be created with a BPD:

- Collaborative (Public) B2B Processes
- Internal (Private) Business Processes Collaborative B2B Processes

A collaborative B2B process depicts the interactions between two or more business entities. The diagrams for these types of processes are generally from a global point of view. That is, they do not take the view of any particular participant, but show the interactions between the participants. The interactions are depicted as a sequence of activities and the message exchange patterns between the participants. The activities for the collaboration participants can be considered the "touch-points" between the participants; thus, the process defines the interactions that are visible to the public for each participant. When looking at the process shown in only one Pool (i.e., for one participant), public process is also called an abstract process. The actual (internal) processes are likely to have more activities and detail than what is shown in the collaborative B2B processes.

Figure 4.28 is repeated in Figure 4.29 to show an example of a collaborative (public) B2B process.

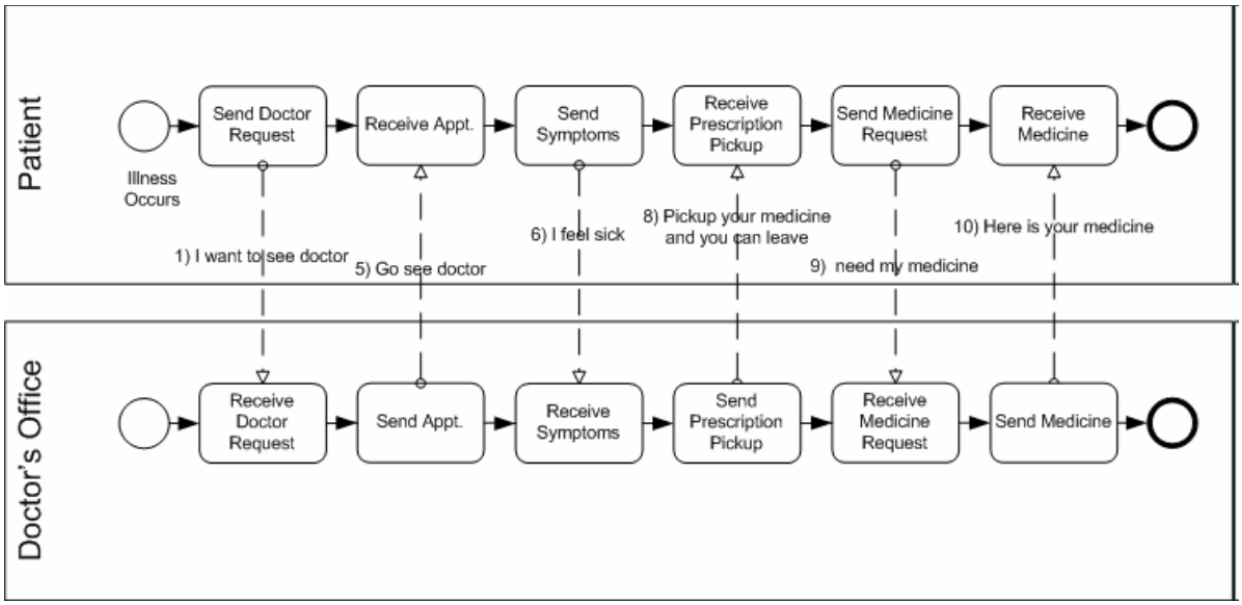


Figure 4.29: An Example of a Collaborative B2B Process

An internal business process will generally focus on the point of view of a single business organization. Although internal processes often show interactions with external participants, they define the activities that are not generally visible to the public and are, therefore, private activities. If swimlanes are used then an internal business process will be contained within a single Pool. The Sequence Flow of the Process is therefore contained within the Pool and cannot cross the boundaries of the Pool. Message Flow can cross the Pool boundary to show the interactions that exist between separate internal business processes. Thus, a single Business Process Diagram may show multiple private business processes.

The modeling of business processes often starts with capturing high-level activities and then drilling down to lower levels of detail within separate diagrams. Figure 4.30 shows an example of a high level process, captured for a BPMN case study, which is basically a series of Sub-Processes with three decision points in the Process.

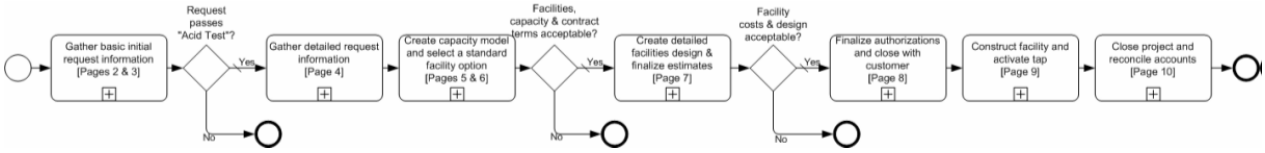


Figure 4.30: High-level Business Process Example

Figure 4.31 shows the details of the first Sub-Process of Figure 4.30. This diagram employs two (2) Pools; one for the customer and one for the company providing the service. Note that this diagram shows the internal business process for the company and shows an abstract process for the customer (i.e., the customer process only includes the activities used for communicating through Message Flow to the company). The activities within the company are partitioned by Lanes to show the departments or roles responsible for their performance (e.g., System Coordinator, Business Development, Legal, and Retail).

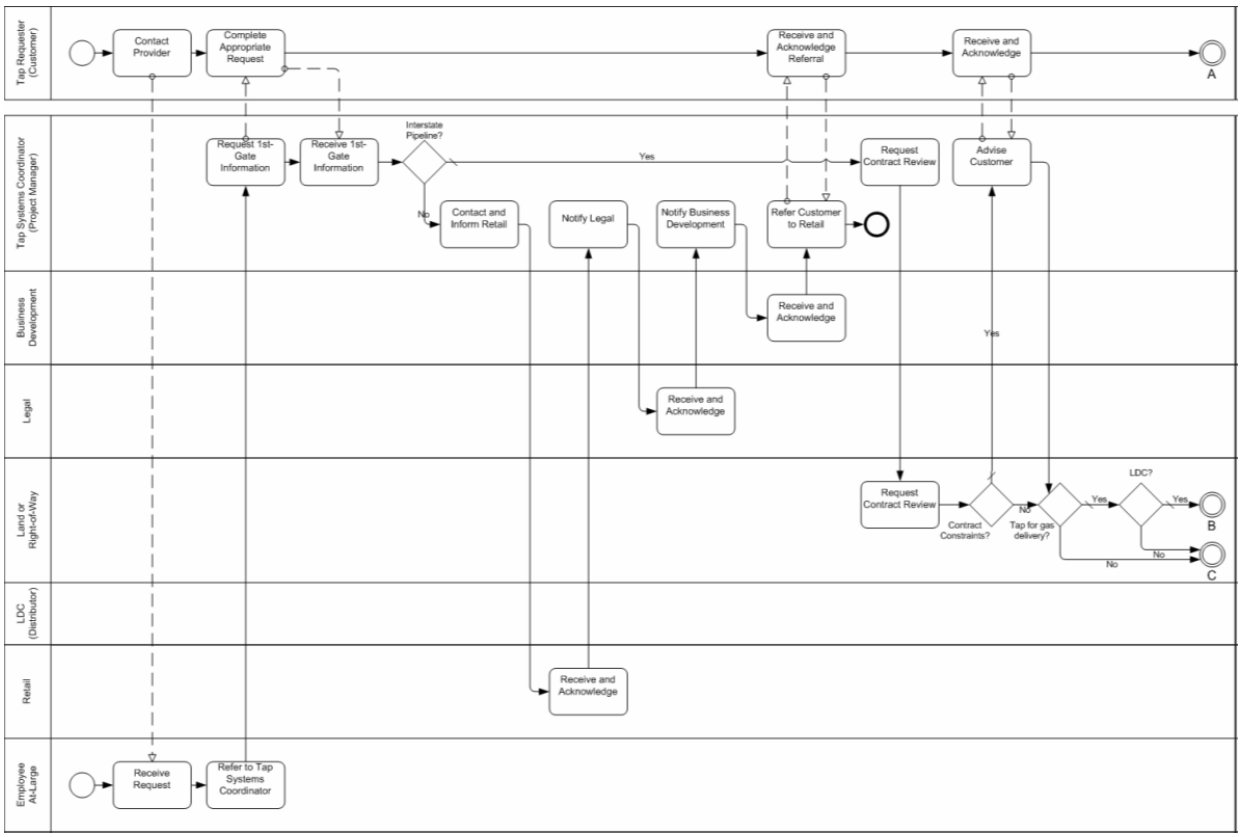


Figure 4.31: Lower-level Business Process for the High-Level Example

4.6 A case of BPMN and application development

In this section we present a simple BPMN teaching case - “Pizzeria Mamma Mia”- that is meant to make students understand how one can start from an informal description of a problem, transform it in a formal specification of the business problem using the BPMN language, and implement this formal business model in a software application that supports and automates the execution of the process.

4.6.1 Informal description

Pizzeria Mamma Mia has problems with the preparation and delivery durations of pizzas. Most clients complain that they have to wait a long time until they receive the pizza they have ordered. Papa Mario Francone (the owner of Pizzeria Mamma Mia) has decided to model and analyze the internal processes in his pizzeria.

The process starts when a client who calls the pizzeria’s telephone number. The employee at the counter takes a note of the order and of the delivery address and gives it to the chef. The chef takes the order and

decides upon the ingredients necessary for the preparation of the ordered type(s) of pizza. Pizzeria Mamma Mia can deliver 25 different kinds of pizza's, which are divided into two categories: standard pizzas and specialties. For a standard pizza all ingredients are available in the kitchen. Ingredients for the special kinds of pizza are kept in a freezer in the storage room, so the chef has to pick them up first.

When all ingredients are available the chef will prepare the pizza and bake it. Unfortunately, the oven is quite old and sometimes the temperature in the oven rises to the level that pizza's get burned. In that case the chef must repeat the whole procedure starting from the selection of ingredients.

If the pizza is ready, the chef gives it to the employee at the counter who puts it in a box and gives it to the delivery guy who takes his motorbike and brings it to the delivery address. He receives the payment and gives it to the employee at the counter when he returns. In turn, the money is put in the cash register. Alternatively, if the client collects the pizza personally, the employee at the counter gives the pizza directly to the client and cashes the money for it.

4.6.2 The challenge

1. Make a BPMN model of the process as described above. The model should reflect clearly which members of the staff are involved by each activity and which systems (motorbike, telephone, oven, pizza boxes, etc.) are required during which activities and are used by which staff members.
2. Change the process such that the order intake will from now on use a software application.

4.6.3 Solution

For the completion of these two assignments we use the software development platform Bizagi. There are other similar commercial products on the market, but Bizagi has the advantage of offering a free fully compliant BPMN modeler. Furthermore, Bizagi also offers a free (for not-for-profit purposes) application development environment, called Bizagi Studio, which we will use for completing the second assignment. This environment allows us to develop a web application (design a data base, the application interface and the application logic) with little technical knowledge and based on the initially created process specification. Download Bizagi Studio from bizagi.com.

Bizagi Studio will take you through all the steps necessary, start at step 1 "Model Process" in the Bizagi Studio menu screen (see Fig. 4.32).

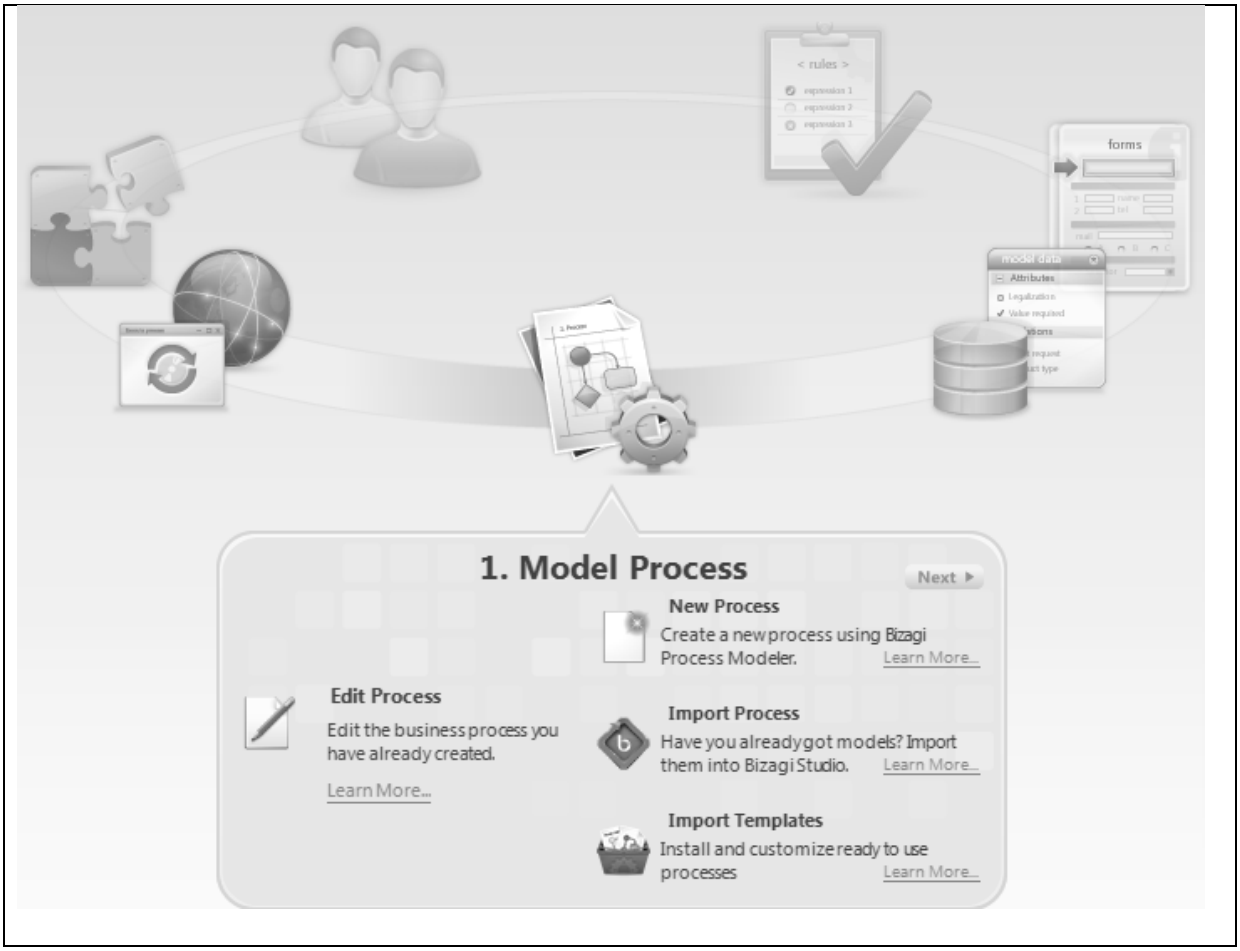


Figure 4.32: Bizagi Studio's "model process" screen

1. The problem described informally in the text above is modeled in this first BIZAGI step as an BPMN model in Figure 4.33.

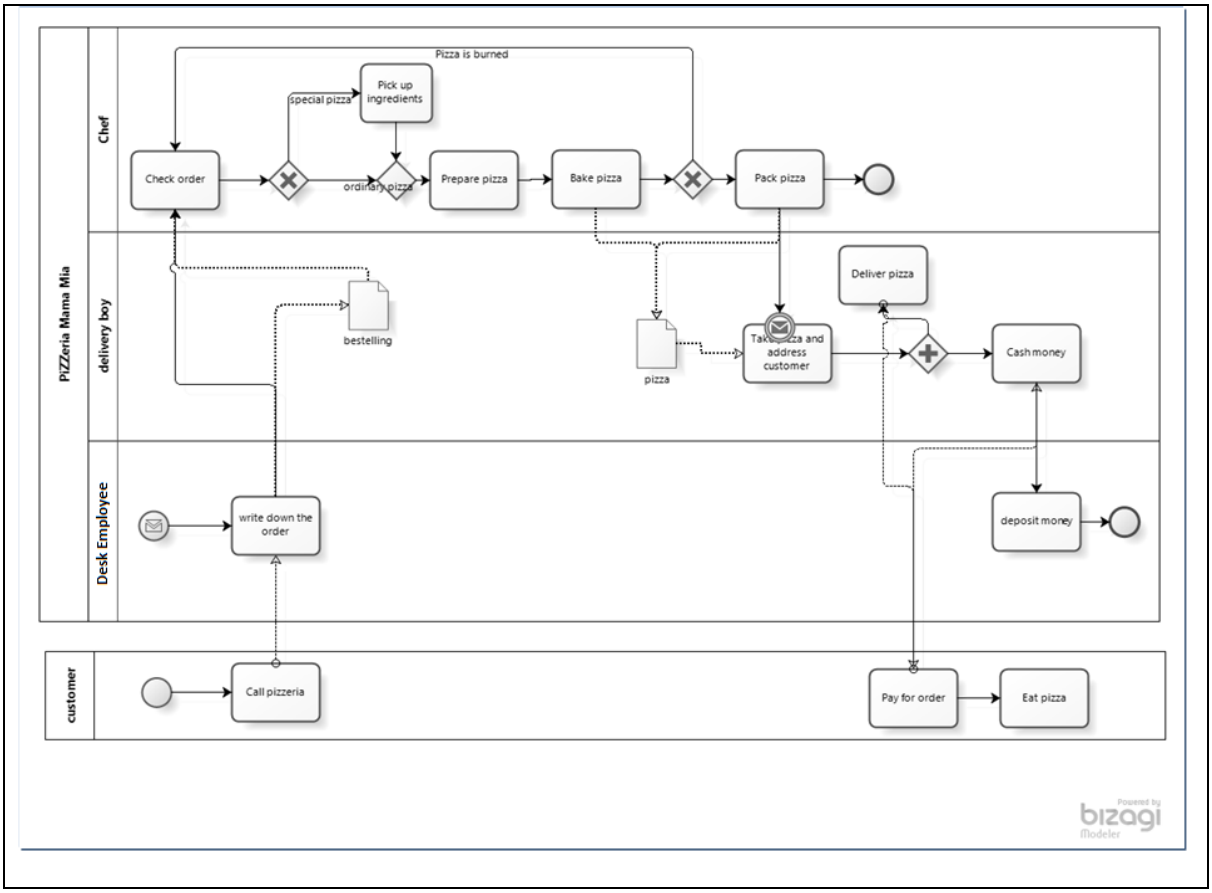


Figure 4.33. Pizzeria Mama Mia's process

2. The modified version of the process is shown in Figure 4.34. The implemented part of the process is shown in the rounded rectangle.

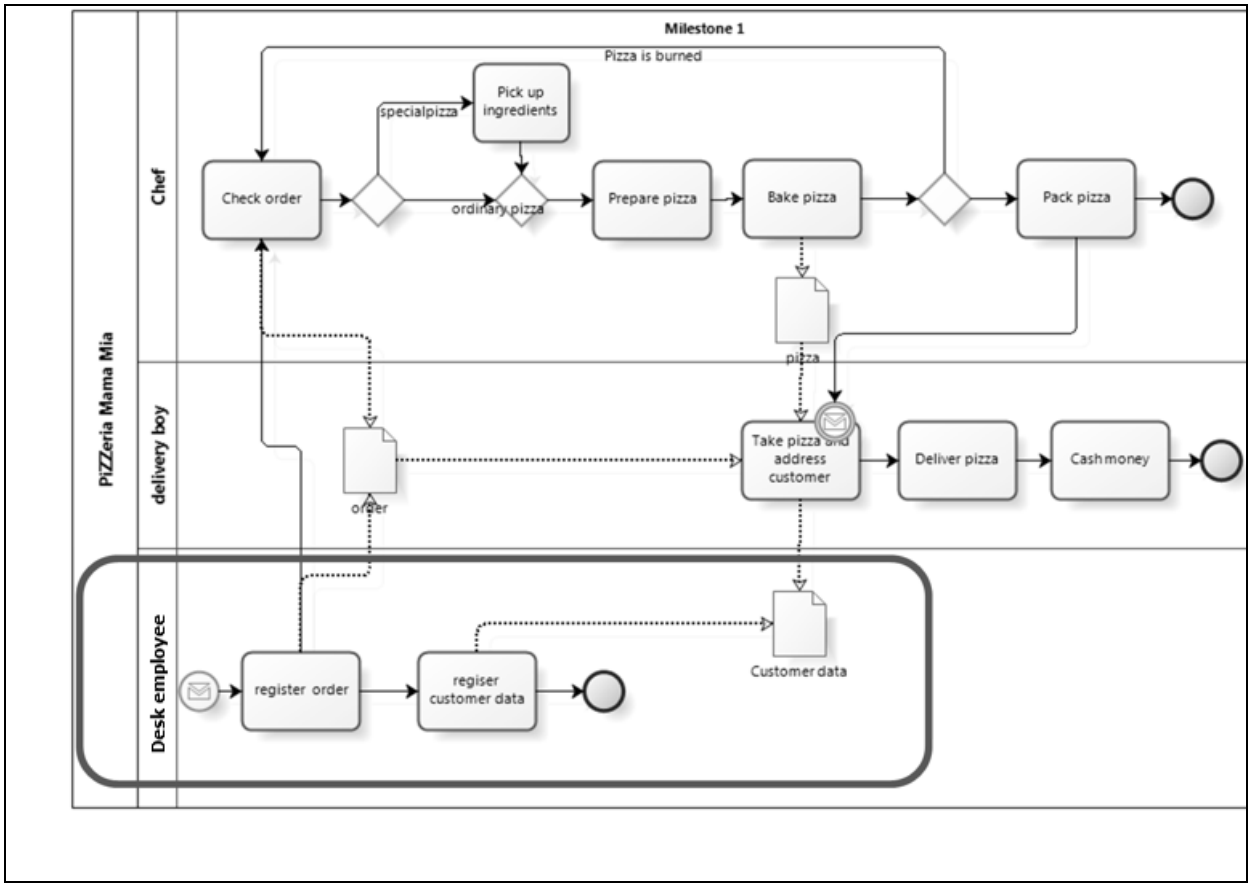


Figure 4.34 The new situation

3. When you have created the process model, continue with step 2 in the modeler: “Model data” that matches the data needed for the process. When you have opened the data modeler, click the “Entity” button to start the entity creation wizard (see Figure 3.35). This wizard lets you fill in the entity properties as a name and a description and continues in the next screen with the attributes. Each attribute can be given a name and a type, default value and length (just as in Microsoft Access). Click finish and you have created your first entity. After the entities have been created, one can define the relationships between these entities. To do this click the “relationship” button to open the relationship wizard. First select the two entities you want to define the relationship for. The next screen, Figure 3.46, shows how to configure the type of relationship (the cardinalities), for example a one-to-one, one-to-many or many-to-many relationship. An example of a completed data model is given in Figure 3.47.

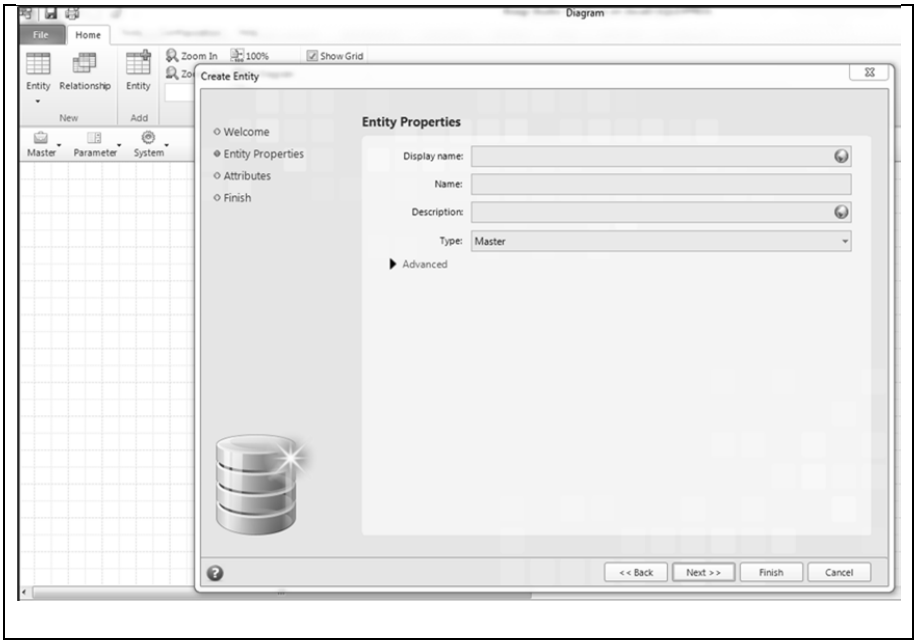


Figure 4.35 Entity wizard

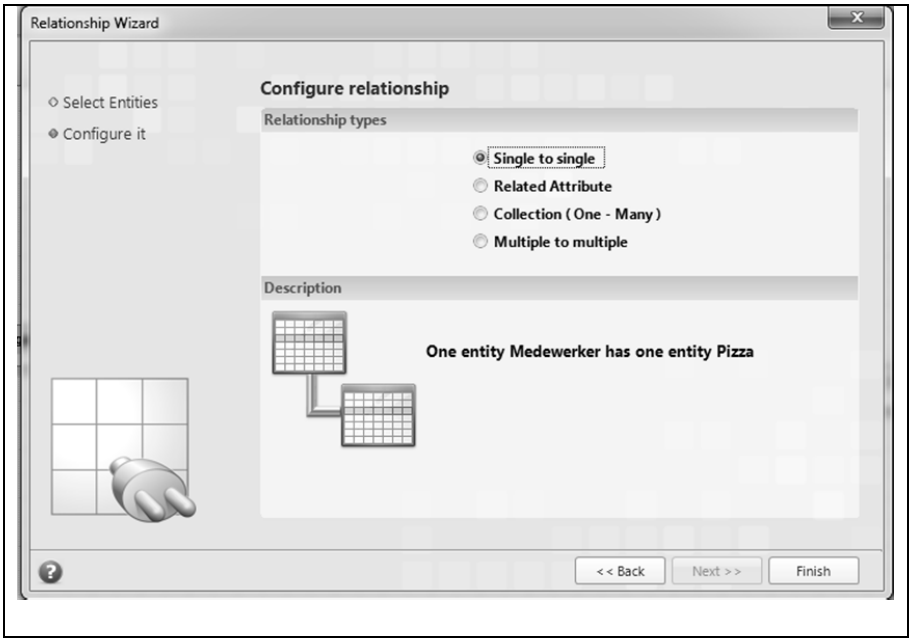


Figure 3.46 Relationship wizard

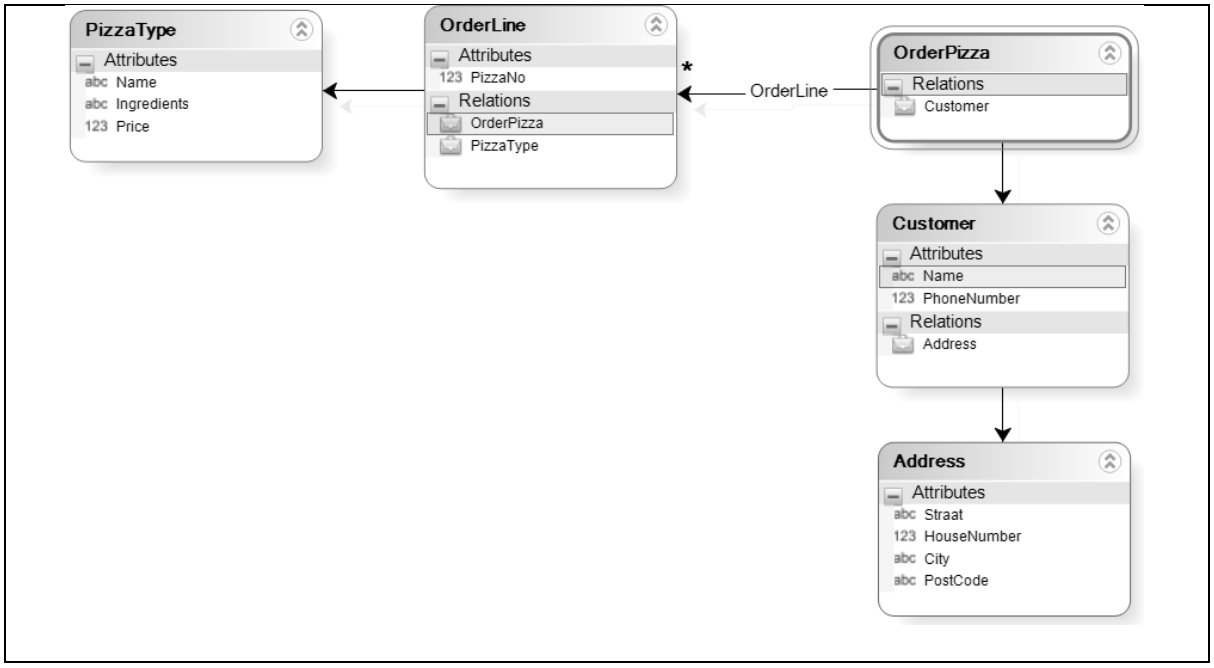


Figure 4.37 Data model in Bizagi Studio

As you can see, there are multiple differences with how a data model is created for MS-Access in chapter 2. The modeling approach for MS-Access is highly relational database focused. When modeling in MS-Access, the user has to think how s/he sees reality and converts this to a model for a database. Bizagi, as a BPMN modeler, is more process centric: the user thinks about steps in the process and models these process steps using the BPMN ontology (categories and relations) and language (related representation tools). Only after modeling the process with this ontology, the users starts to think about a data model. Designing the data model is easier in Bizagi (excellent representation, I would say) as the software helps the user with creating the data model by wizards, but of the world does not consist of processes or one needs to model processes differently (J. Recker, 2010), Bizagi is less helpful.

4. The process model is used as a basis to create an interface design for each step. Use the "create forms" option to open the form design interface. See Figure 4.38 for an example.

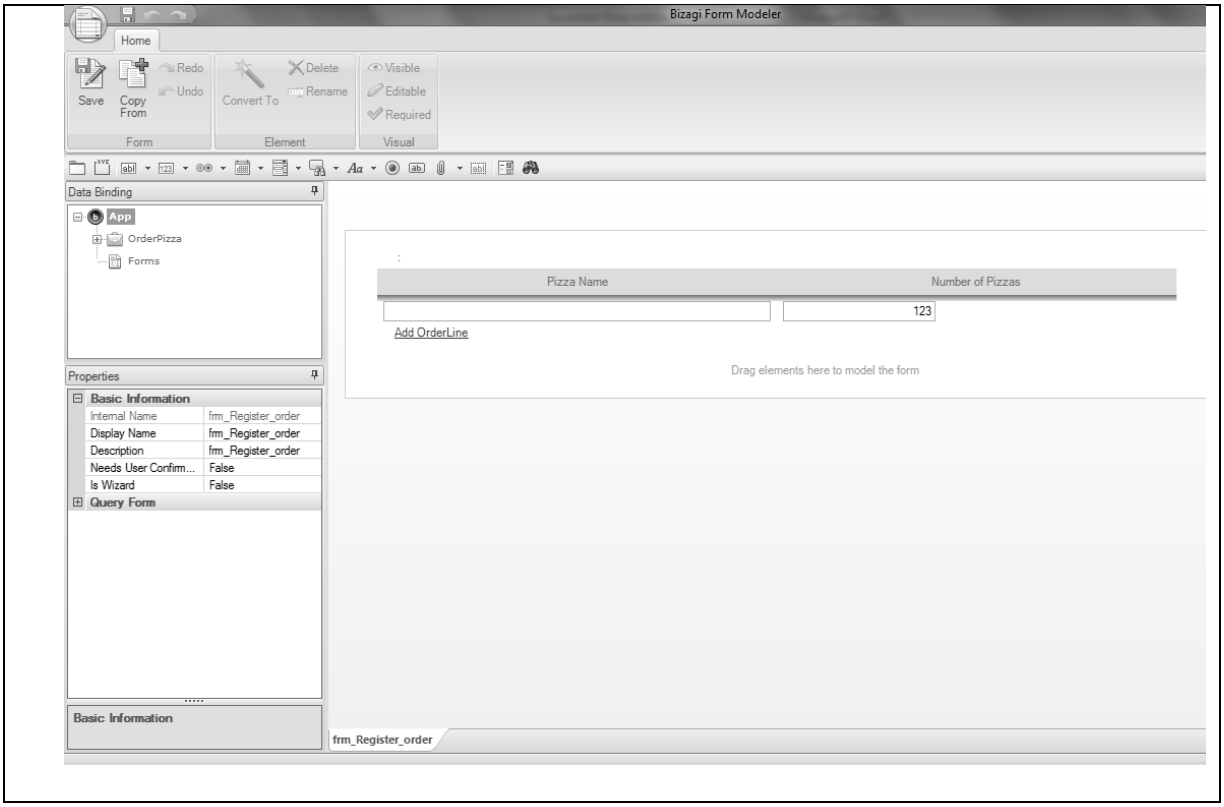


Figure 4.38. Example of Interface design for the "register order" step

5. In Figure 4.39 you see the finalized screen of the “register order” step when executing the process.

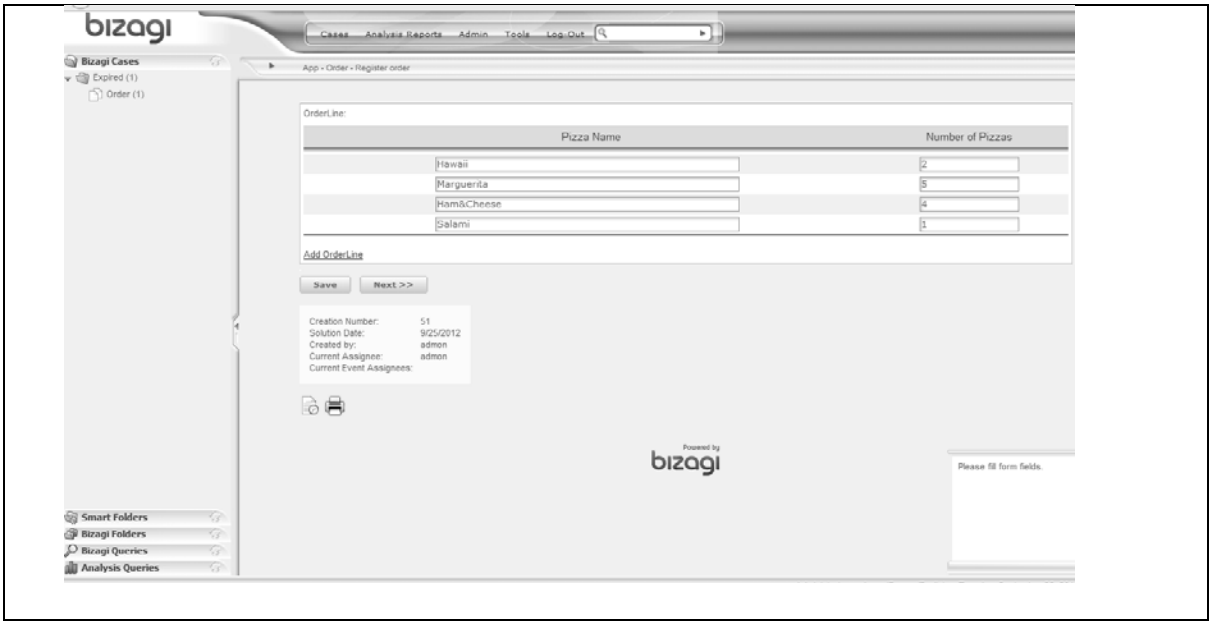


Figure 4.39 Screenshot of the execution of the "register order" step

4.7 Further study

The computer science discipline has developed various modeling tools in recent decades, which facilitated to systematically describe an organization, diagnose and define how information systems can contribute to a better process. In this context, one of the most cited books is of Scheer (Scheer, 2000) where Scheer presents the ARIS method, which is owned IDS which performs studies for the implementation of SAP (the world leader in business software). For a discussion on the usefulness of modeling languages see Recker et al (J. C. Recker, Rosemann, Indulska, & Green, 2009). For more extensive study of workflow management and business process management also see Van der Aalst and Van Hee (Van Der Aalst & Van Hee, 2004). For more on BPMN and Bizagi go to www.bpmn.or and www.Bizagi.com.

4.8 Exercises

1. SmileYou. Mary-Ann's fashion company contracts production to producers in China, Turkey and India. For this procurement managers specify products, negotiate and contract with producers. Each season volumes of end products will be delivered to SmileYou warehouses in Western Europe, Northern America and the Middle East. In 2010 over 30 million euros of production orders were released. After some checking of what finally entered the shops, there is an estimate that almost 10% of all the products never arrived at the shops; a loss of 3 million. Consequently, Mary-Ann wants to have the logistic processes from manufacturer to shop carefully documented,

and wants concrete solutions of the problem by organizational measures, process designs, information systems and data where possible.

   a. How can this be done in an informal way by MS-Visio?
   b. How can this be done by BPMN?
   c. What are differences, advantages and disadvantages of both methods?

2. Student administration. Describe the organizational measures, knowledge needed, processes, information and systems that are needed to select applicants for 10 different university bachelor programs. What will be needed for an average university (about 25000 students, 5000 applicants per year) to process candidates within 2 weeks? Also create the BPMN model and realize the process in Bizagi.
3. Three questions for reflection:
   a. What is the relation between an ontology as a philosophical understanding of the essence of things and a priori's and modeling languages?
   b. Why is knowledge based on a priori's not empirical knowledge?
   c. How can it become empirical knowledge and what is the role of a posteriori's here?
4. Final question: How can one enable the collective brain by the Kantian inquiring system and related tools? So what information or knowledge are we talking about and how can this become collective in an organizational setting?

5. The Hegelian perspective and information triangulation

5.1 Hegel and critical information analysis

The Hegelian inquiring system suggests that information is mainly subjective and contextual, and that each statement (thesis) may be countered by an antithesis on the same or different evidence (Churchman, 1971; Wijnhoven, 2012). The major job for people is to know the evidence on which each (anti-)thesis is based, criticize it and develop a plausible synthesis on top of this reasoning and evidence. For information management this means that there is data which is used to support a certain view on reality, an antithesis which reflects a counterview, and a synthesis which combines the two to a new understanding of reality. This has major consequences for Internet information, which mostly lacks control over the truth (Wijnhoven, 2012). Think for example of any person who would want to communicate negative views about Grand Café. He could tell on the internet (via social software) about his "experience" and exaggerate some minor weak points (this is named the thesis). If anybody reads this review s/he may reconsider about going to Grand Café. It is important to check in this case who has been the author of the review and what has been his or her viewpoint. Is it a complete picture of the dinner or does s/he only talk about the negative things? Of course, the owner of Grand Café could respond and state how terrific they are (and thus states an anti-thesis) but who should the reader believe and trust? Only by asking the right kind of critical questions an information consumer –following Churchman (Churchman, 1971) and Wijnhoven (Wijnhoven, 2012) named "information slave"- can find out that maybe the thesis or anti-thesis is incorrect and starts believing one of them or better develop an opinion of his or her own. Often, however, both thesis and anti-thesis are hard to reject and have considerable good evidence in support. In that case, the information slave must emancipate and build his or her own opinion, i.e. synthesis.

Information is never necessarily complete and may be biased towards the support of a particular view. This is obviously so, because reporters and writers can never be complete and are selective to serve the particular interest of their funding organizations (mostly the newspaper shareholders, the real information masters) and their information consumers. Alternatively, information consumers as information slaves, mostly lack the tools, time and means to critically check the veracity of the information that is presented, but some basic insights and skills to do so can help a lot. Therefore we present the critical information triangulation method later in this chapter.

Although Internet information is mostly biased and unreliable, one may ask if officially well-developed information systems in organizations are much better. The answer to this question is often yes, but also often no, because the realization of reliable information by systems is very complex as well. Many information systems are developed to deliver management information. The main difference with Internet

information is that there is mostly no counter evidence in an information systems context, making it actually more sensitive to sources of unreliability and confirmation bias (Fischer et al., 2011). Management information must of course represent correct measures of business performance and should help management to make the right payment and investment decisions. This sounds simple but is difficult to realize in an objective and non-biased way. The biases may be caused by problems with five key parameters of management information (Hofstede, 1981).

1. It is mostly impossible to know all the effects of certain actions on the success of an organization. Often some effects and side-effects are overlooked, thus even if we have correct data about some effects, we may have incomplete views of reality and thus the data are not helping much in decision making.
2. Many actions that managers take are non-routine because managerial decision situations are often unique. Consequently, no routine "if then" rules can be easily found and applied. Experts may be able to handle complex less routine cases, but the related expertise is often not available. So even if we know all the data about everything, the lack of causal rules will disable us to make optimal decisions.
3. When no knowledge exists about the possible effects of certain actions (e.g. if a payment raise will result in cost covering productivity increases) often decisions have to be made by trial and error (when these decisions can be reversed!) or by pure intuition. So called "evidence" in these situations may be more political and subjective, less objective and just an argument to legitimize what is in the interest of the information master.
4. Management information systems cannot easily help on situations where the objectives are unclear and when we do not know the impact of actions. For example, one may be able to measure less sales, but if we do not know for example how much a larger sales department and larger advertising budgets increase sales and profits, it is difficult to interpret sales data for deciding on the best department and budget size. One even may state that a sales department's performance should not be primarily measured by sales volumes because exposure in the media, networking and market knowledge may be more important for decision making. Output measures can be very ambiguous in this context.
5. Output measurement is also sometimes difficult and may also result in the wrong incentives in organizations. For example, what is a good output measure for universities, the number of diplomas granted or realizing the best fit between a person and the labor market? Number of diplomas maybe easy to measure, but this could result in an incentive to send many poorly educated people on the labor market. Fit to the labor market may be better, but how to measure

this? If fit to the labor market cannot be directly measured, we may try to find surrogate measures by sending out questionnaires to alumni (but what to ask them?) and finding data about the employment or unemployment numbers of alumni (but do employed alumni have jobs for which they have been trained?). When these ambiguities cannot be solved, decision making will become politics and information only serves the legitimization of decisions.

6. However, when reliable disambiguous data can be gained, and objectives are clear and the relation between actions and impacts are known, information can be used for decision optimization. It can also become routine decision optimization when the decisions happen frequently in the same context, so that it is worthwhile to develop in routines and related automated business processes.

These issues of objectives, measures, effects and repetition identify each a type of control, with a different role for information (See Figure 5.1).

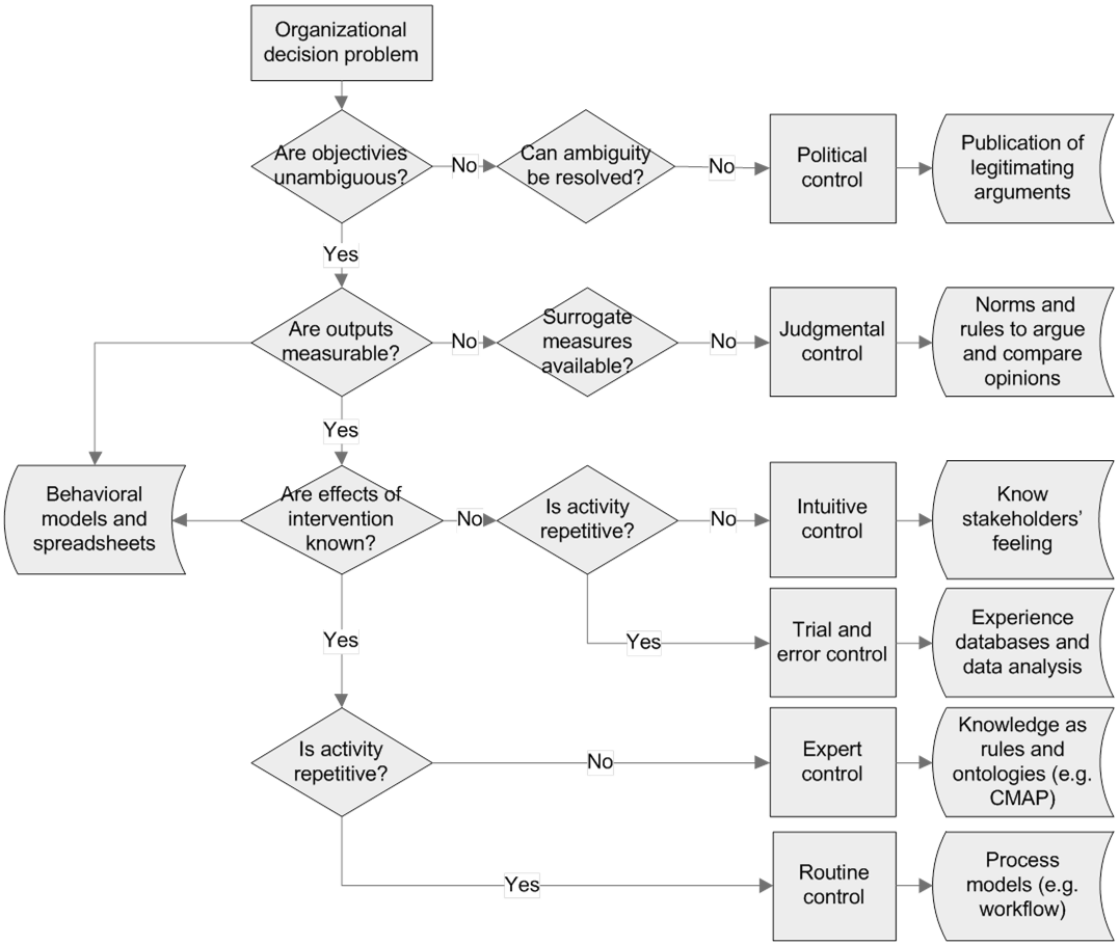


Figure 5.1: Types of control and the role of information presented by . Source: (Hofstede, 1981).

5.2 Goal of this chapter

Obviously, information is not neutral according to the Hegelian inquiring system, and its role is different depending on the decisional context. Of course, we want to avoid information errors on the Internet and in information systems contexts. This is realized by low versus high content and publishing control respectively. However, full control is impossible and errors are unavoidable, and consequently, we will focus in this chapter on the information slave, and emancipating this slave by giving him or her the right kind of key critical information triangulation tools for both low and high content creation control contexts, i.e., the Internet and information systems. An important part of this method is dialectics, i.e., the search for anti-thesis (a fully contrasting perspective on an opinion already held) and the development of synthesis that integrates dialectically opposite views by a higher level insight. This method has been used for strategic decision making frequently (Schwenk, 1990), but is useful for any complex decision making that has to cope with multiple interests. Additionally, we give insights in a method named brand monitoring and sentiment mining, which totally changes the Hegelian perspective from one person who has no control over what others are saying about him or her. This "slave of the crowd" perspective is given to systematically checking what other says, aggregating it on one place and analyzing it regarding positive or negative statements.

We assess the value of information through critical Internet information triangulation in section 5.3, information systems information triangulation (also named assumptions checking in (Wijnhoven, 2009b)) in section 5.4, and aggregating and mining the opinions expressed by others in section 5.5. For this we apply a variant of the general information management model as described in chapter 1, which starts with scoping, finishes with actual understanding of the world. See Figure 5.2.

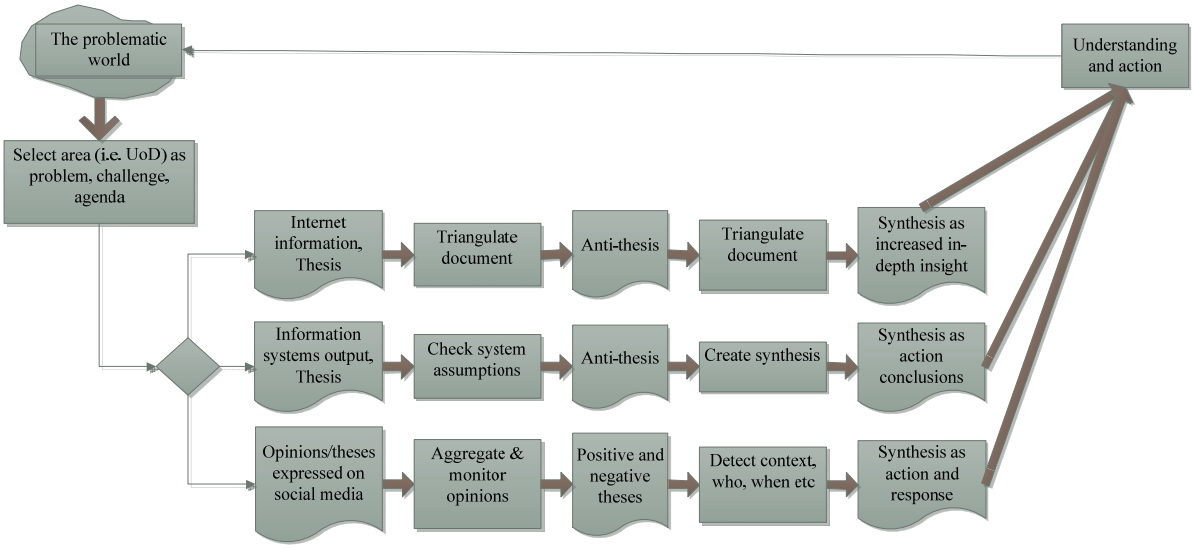


Figure 5.2: The Hegelian information management model.

Ultimate goal: Assessing the value of information through information triangulation, information systems assumptions checking, and opinion mining.

5.3 Internet information triangulation

Triangulation is about information values. There are four kinds of triangulation, which Wijnhoven took from Denzin (Denzin, 2009; Wijnhoven, 2012).

- Data triangulation focuses on the verification of data which is acknowledged in the source. An example is the 2007 climate change report from the IPCC (Intergovernmental Panel for Climate Change), which contain a few errors which were discovered later. These mistakes were discovered by comparing the data in the IPCC report with the data in other sources.
- Investigator triangulation focuses on the author. In the IPCC report (Wijnhoven, 2012), three commissions are housed. Commission 1 only reviews pure scientific and licensed publications for inclusing in the IPCC report. Commission 2 only points at everything which is already known and written about the climate. Commission 3 focuses on the impact of observations of the climate by using predictive models. Because of this, the report is filled with scientific and non-scientific data. So it is essential to distinguish what is said by Commission 1, 2, and 3. Who says what and what is the motivation of somebody to say something are questions you have to ask with investigator triangulation.

- Theoretical triangulation. You can see the IPCC report in different ways. A skeptical opinion is to think that the report is only in benefit of environmentalists. An economic concept is to see the possibilities of climate change for the development of previously infertile or not exploitable soil. IPCC information can also be regarded from a political scientist perspective by explaining its publications and consequences from an energy geo political perspective.
- Methodological triangulation is divided into four species of data used to develop evidence:
  A. Empirical data in the IPCC reports is for example the evidence that states that glaciers are melting and getting shorter, with measures of sizes and speed of melting. Of course, these measures must be reliable and precise to a certain minimum level.
  B. Interpretive data are data that explain why people do what they do and for example why it is so difficult to change human behavior regarding energy consumption.
  C. Historical data. Although thermometers are relatively new human inventions, temperatures can also be estimated by the size of tree rings. An important question is whether the current increases are more abnormal and extreme than the historical increases or if speed and frequency of change is different.
  D. Critical data. This is data on the impact of IPCC data on society and human behavior. Critical data may show for example that global warming has disastrous effects on the abilities of whole populations, and the data may show the cause being human or not. Well if the cause is human, people have to change and the political consequences will be huge. If the cause is not human, people may want to continue living the way they always did and just wait until a disaster happens (or hopefully not). This is exactly the debate between climate doom thinkers and climate skeptics or between the interests of the environmental movement and the car industry.

Wijnhoven (Wijnhoven, 2012) has the following steps in dialectic triangulation of evidence (see Figure 5.3).

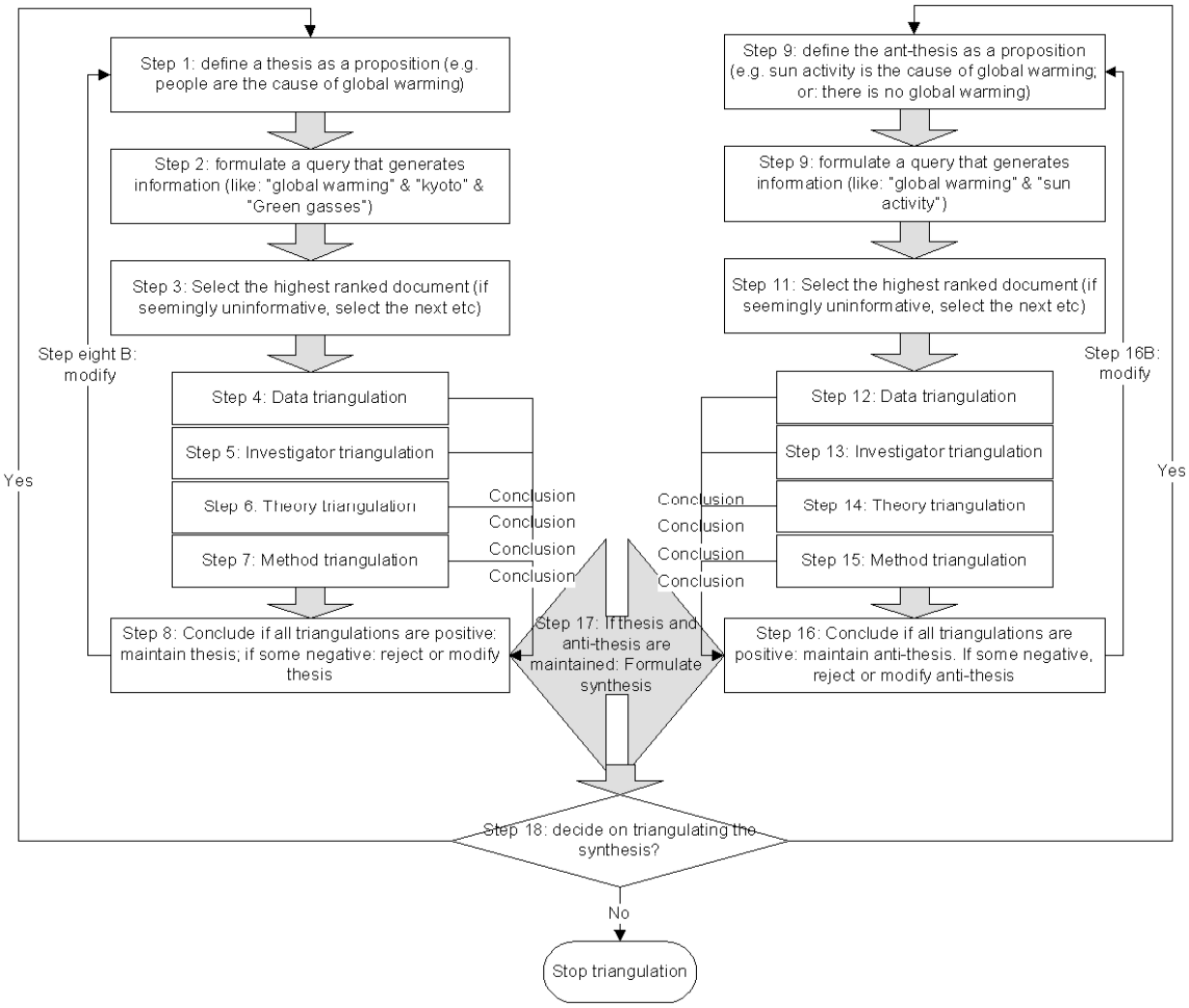


Figure 5.3: The Hegelian information triangulation workflow. Source: (Wijnhoven, 2012)

The aim of dialectic information triangulation is to be critical, but constructive as well by aiming at higher levels of understanding and avoiding cynicism regarding human possibilities to achieve more depth and awareness of problems (Schwenk, 1990). One may for example want to triangulate the thesis "Greece should exit the Eurozone", by searching a document (for example http://www.euinside.eu/en/news/eeag-greece-to-exit-from-the-euro-area-and-to-devaluate-the-drachma) by:

1. Checking the quality of the data that supports the article (i.e. data triangulation).
2. Checking the author and his or her reputation and affiliation (i.e. investigator triangulation).

3. Checking the basic theories, approach, goals-means and cause-effect relations behind the article (i.e. theory triangulation), and finally,
4. Checking the method by which the paper came to its conclusions, like based on empirical research, experiments, interviews, normative reasoning etc. (i.e. method triangulation).

To give this a dialectic turn, an anti-thesis (e.g. "Greece should stay in the Eurozone") can be used to search for alternative articles, triangulate these, and try to draw a synthesis on basis of the triangulation results (e.g. Greece should stay in, rich EU countries will have to take most of the losses, but the Greece financial systems has be revised and controlled by the Euro Bank").

5.4 Systems information triangulation and assumption detection

The same questions as with Internet information can be raised regarding information systems output. To explain this, let's give the fictive example of the triangulation of a management report.

- Context: The annual report of 2009 in company X reports a poor sales results for that year. The Top management shows concern and holds as THEORY (thesis) that the Sales department is not well working, inefficient, and that a solution (intervention) of the problem is needed by cutting down its size by 20%.
- The Sales department head holds as THEORY (anti-thesis) that they have been working fine, but the economy is down, a new competitor took market size and so they need more support (intervention) to beat the competition and be prepared when the economy is better again.

For both Thesis and Anti-Thesis, data, investigator and method triangulation is possible to increase insight in the situation. (Note that the detection of Thesis and Anti-thesis itself is already theory triangulation).

Thesis triangulation:

- Data for inefficiency of sales department (output measurement), which needs an agreement on objectives and measures like sales versus cost. Achieving a high Sales versus costs rate can be recognized as a very short term view on a Sales department's objectives.
- Investigator: The company's controller has an assignment from the top manager to create the data and cannot act independent of this Top management assignment.
- Method: The observations only consist of "Registered sales figures" and "department costs figures" included in the "SAP database".

Anti-thesis triangulation:

- Data for success (number of new clients, feedback on promotion & brand name, number of sales, market-size figures (Nielsen), macro-economic census data. The Anti-thesis thus includes a much richer and especially a longer term perspective of the Sales department's objectives. Objectives definition: sales/costs is too restricted short term for the moment.
- Investigator: The Sales Department head has asked a market consultant to help on the development of a strategy (anti-thesis! In this case), collect alternative evidence and analyze these to make a business case to support the head of department's view and interest.
- Method: The alternative data that are collected consist of market & macro-economic figures, feedback forms and website statements after promotions and fairs. The alternative analyses consist of scenarios of market growth and sales capacity needed. This includes a much richer set of data, aiming at longer term views of the sales department.

Well, how to create a synthesis? First the synthesis may be split in a short term and longer term perspective. If the Top management decides in favor of a longer term perspective, they give in, and decide to cut the Sales department only by 5% (also to show other stakeholders that they are concerned regarding the performance). This policy will be reviewed after 6 months after which further cuts could be taken. The Head of department accepted this offer for the time being, but also acknowledges that of the anti-thesis is incorrect she will probably will have to resign and the department will be fully reorganized.

5.5 Triangulation by brand monitoring and opinion mining

How frequently have you debated perceptions of "the general public" or a target group like "the client" without any evidence except of a few ad hoc meetings with some representative people? Such discussions can easily become impossible to finish, unless someone has done a systematic market research or someone is able to decide because of his or her seniority, expertise or higher level authority and responsibility. Market research (an empirical solution!) is often too slow and too expensive to be used all times and deciding by authority can, according to the Hegelian ambitions, easily result in less than good decisions. To liberate the information slaves, brand monitoring and sentiment analysis can be very useful here, because it is cheap, fast, and makes optimal use of opinions of the masses. Nowadays many views on products and services are delivered on social media platforms like Twitter, Facebook and blogs whose content can be collected by brand monitoring.

Dan Schawbel gives a set of key activities one can do to aggregate opinions (Schawbel, 2009). He states that "….brand monitoring has become an essential task for any individual or corporation. Years ago, when people talked about our brands, it was behind our backs and we almost never found out about it.

Today, most of these dialogues are right in front of our own eyes and the number of locations where our brands may be cited is astronomical!"

We must keep in mind that conversations about our brand are being held on the web with and without our consent. One can choose whether to be a participant or an observer or outcast. Before selecting a passive approach like the observer or outcast, be reminded that these conversations are not necessarily positive. Besides this, the speed of these messages on the internet is also very high. Something that starts out small, can turn into national news in a timeframe of several hours.

Here are free tools suggested by Schawbell, but classified slightly differently here.

1. Feedreaders. Feedreaders like Google reader are easy for sorting feeds, bookmark/favorite them and share (give value) them with your network. One can also register for a Delicious account, which can help you sort and organize blogs that mention your brand. Once you have set up these two accounts, the following tools will help locating articles that mention your brand, feed them right into Google reader and allow you to manage them by Delicious. Google Alerts are email updates of the latest relevant Google results based on your choice of query or topic. You can subscribe to each alert through email and RSS. The alerts track blog posts, news articles, videos and even groups.
2. Blog post monitors. Technorati is the largest blog search engine in the world. Technorati tracks "blog reactions," or blogs that link to yours. Search for your brand on Technorati, and subscribe to RSS alerts so that when someone blogs about you, you find out. Blog comments. Backtype is a tool for monitoring blog comments by letting you find, follow, and share comments from across the web. Whenever you write a comment with a link to your Web site, Backtype attributes it to you. Use it to remind yourself where you commented, discover influencers who are commenting on blogs that you should be reading, and continue conversations that you started previously. You can even subscribe to these comments using RSS. coComment is another tool that will help you manage your comments across the web.
3. Social comments. Yacktrack lets you search for comments on your content from various sources, such as Blogger, Digg, FriendFeed, Stumbleupon, and WordPress blogs. For instance, if you comment on a blog, you can locate other people who are commenting on that same blog post and rejoin the conversation. "Chatter" tab allows you to perform keyword searches on social media sites and then notifies you of instances of your brand name. Yacktrack's search page results also

give you an RSS feed for the search term. You can also use Commentful and co.mments to track your social comments on the web.

4. Discussion boards. Along with blogs and traditional news stories, discussion boards are channels where people can gather in a community and talk about you. Most people disregard discussion boards until they see other sites commenting on information viewed on them. Boardtracker.com gives instant alerts from threads citing your name.
5. Microblog search. Using Twitter search, you can locate any instances of your name and decide whether you want to tweet back or ignore them. It really depends on the context and meaning of the tweet. Conduct a search for your name, your company's name, or various topics you're interested in and then subscribe via RSS. Twilert and TweetBeep are additional tools you can use to receive email alerts.
6. Social account aggregators. Friendstream is a social account aggregator, which gives the ability to take all of your social accounts, such as YouTube, Delicious, Twitter, blog, and Flickr, and pull them together into a single (Friend) feed. You can conduct searches on your brand throughout all social networks at once using this search engine. Aside from learning about the latest video or tweet related to your topic, you can analyze comments that people make under them. FriendFeed users tend to favorite and comment on what you share and tracking it will become more important as this service grows in population. You can also receive alerts straight to your desktop with Alert Thingy.
7. Social search. Social Mention is a social media search engine that searches user-generated content such as blogs, comments, bookmarks, events, news, videos, and microblogging services. It allows you to track mentions of your brand across all of these areas. The results are aggregated from the top social media sources, such as Flickr, YouTube, Digg, Delicious, Twitter and more. Like the other services, you can subscribe to your results by RSS or email. Other social search engines include Serph and Keotag.
8. Interactive search. While all the other tools listed are quite rudimentary, this one is rather complex and intelligent. Instead of being hit with hundreds or even a thousand results for your brand name, Filtrbox only delivers the most relevant, credible mentions of things you need to track. Its "FiltrRank" technology scores content based on three dimensions: contextual relevance, popularity and feedback. You can look back to previous searches 15 days out for free as well.

So, if you have collected a lot of opinion statements in this way, how can you turn this in a sentiment report? For this, you first need to know what an opinion is and next how reports can look like. An excellent resource on opinion mining is from Bing Liu (Liu, 2011), on which I base most of the

description below. There are also several webservices for creating sentiment reports. Two examples of these are the website Sprout Social (sproutsocial.com) and Opinion Crawl (opinioncrawl.com).

Opinion mining is the analysis of opinions expressed on social media. Synonyms are for example sentiment analysis, sentiment mining, subjectivity analysis, affect analysis, emotion detection, and opinion spam detection. An opinion is a sentiment expression and can be of a regular or comparative kind.

1. A regular opinion expression states an opinion on a specific target subject, like cars, phones or government services, which can be a :
    a. Direct opinion statement like “The touch screen is really cool” or,
    b. An Indirect opinion statement like “After taking the drug, my pain has gone”.
2. A comparative opinion expression compares the values of multiple entities, like “iPhone is better than Blackberry”.

In the further explanation of opinion mining we focus here on regular opinion expressions.

One can look at a sentiment expression at three levels:

1. Document level, i.e., does this review represent a positive or negative sentiment?
2. Sentence level, i.e. is each sentence positive or negative?
3. Entity and aspect level, i.e. an evaluation of parts of an entity or the full entity for one or multiple aspects. More precisely, an entity like a mobile phone, can have multiple components (like a screen, memory, sound, light), which each have certain attributes (respectively size, size, decibel, and strength for the four mentioned as components). Components and attributes are both named “aspects” or “features”.

Any opinion expression includes a target (a specified entity and related features and aspects), a sentiment (which can be positive, negative or neutral), a person who holds the opinion, and a moment/time when the opinion is expressed. The last two, we also name the context of the expression. Lui (Liu, 2010) formally describes an opinion as a quintuple. An opinion is a quintuple e_j, a_{jk}, so_{ijkl}, h_i, $t_{l,}$ where e_j is a target entity or object of analysis, a_{jk} is an aspect/feature of the entity $e_{j,}$ and so_{ijkl} is the sentiment value of the opinion from the opinion holder h_i on feature a_{jk} of entity e_j at time t_l. Sentiment analysis aims at reports that summarize opinions for a product or service from multiple sources or opinion holders. For this many different reports can be produced for example per aspect/feature, per target, and per time. Important question is of course if the right kind of features (a priori) and products (synthetic a priori’s) have been analyze and compared.

Sentiment analysis aims at reports that summarize opinions for a product or service from multiple sources or opinion holders. For this many different reports can be produced that are summarized in Figures 5.3, 5.4 and 5.5.

Figure 5.3: Output of twitrratr when searching for product "st ives"

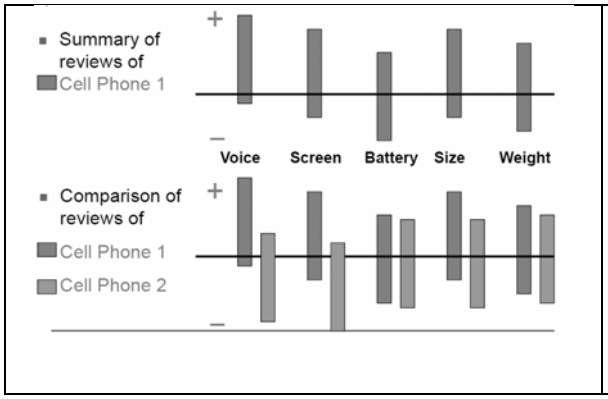


Figure 5.4: Example opinion observer for a cell phone (product name anonymized)

Figure 5.5: Opinion trends for cell phone of Figure 5.4

A web service which converts any text document to a sentiment report is www.alchemyapi.com. Also, at http://sentistrength.wlv.ac.uk/ you can try some sentiment analysis. Figure 5.6 shows an example of a search by socialmention.com. You can see the *strength* of the brand name, which is the likelihood that the brand is discussed on social media. The *sentiment* is the ratio of positive comments over the negative

comments. The *passion* is the likelihood that a person who talks about the brand once will do so repeatedly. And the *reach* is the number of unique authors. As this is a free tool, you cannot expect it to be infallible, but it does give a good idea of what sentiment mining can give.

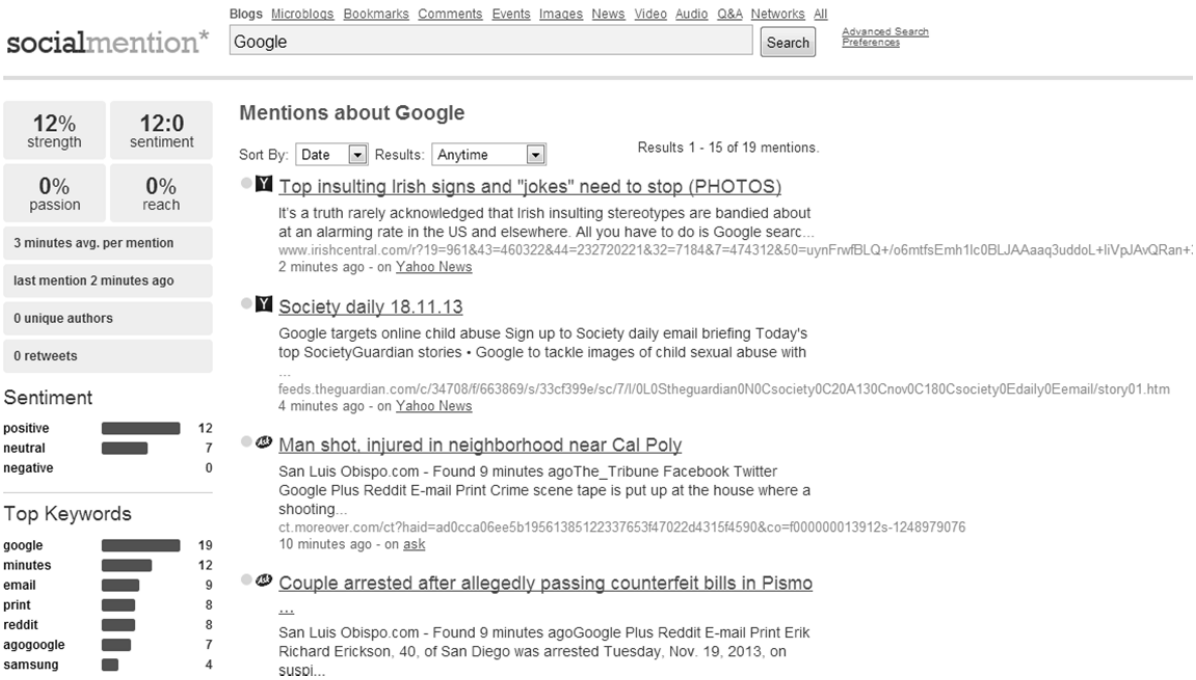


Figure 5.6: Search on SocialMention

Social media triangulation can be a very useful way to gain access to a large amount of information. By combining data mining in social media with triangulation techniques you can gather the data and check its reliability. As one cannot really expect to find scientific theories and methods on social media, method and theory triangulation will probably not yield any interesting results. Data- and investigator triangulation however can make good use of social media. Since a lot of information about individuals is available on the social media, it is possible to use this as a source to determine the credibility of an author (investigator).

5.6 Further study

See Wijnhoven (Wijnhoven, 2012) on Hegelian information triangulation, Hofstede (Hofstede, 1981) on systems assumption checking and Schwenk (Schwenk, 1990)on dialectics in decision making. More details on sentiment mining can be found in Liu (Liu, 2011) and on the website of AlchemyAPI.

5.7 Exercises

1. Company X was advised after a short informal visit of an organization consultant to outsource its financial affairs to a Consulting firm. The Consultancy firm has an excellent demonstration of their capabilities in processing all administrations (human resources, pay roll, client transactions, and internal bookkeeping and control). How can the Company triangulate the Consultant's offering?
2. Which way of control of the assumptions is the best if….
   - You were at a service desk of a computer shop, where objectives are unambiguous, outputs are not measurable and there are no acceptable surrogate measures? What sort of control would you use?
   - A big soda company is implementing a new system, where objectives are unambiguous, outputs are measurable, effects of the intervention are known, but the activity is not repetitive. What sort of control would be the most suitable here?
   - A big machine park is using machines: outputs are measurable, effects of an intervention are known and the activity is repetitive. What sort of control is the most suitable in this situation?
3. Catherine Malone is the head of the sales department of SmileYou. She is responsible for 10 sales persons that regularly visit companies and national headquarters of SmileYou, but which also does all the PR and even participates in product design and product decision at SmileYou. Catherine has a base income of 30,000 euro per year and an additional payment that can go up to 9,000 depending on the net profit of the firm. This year was not very well, and Catherine was not offered a salary supplement. Additionally, the department's costs will have to be reduced by 20% (two people out), although the workload will remain the same. Catherine is furious, because the sales department is working fine. The economy is just down and there is a new competitor. The sales department needs more money to invest and beat the competitor so they are prepared when the economy is better again.

   For this assignment, you have to:

   1. Define a thesis regarding this issue (Management) with intervention mentioned too.
   2. Define an anti-thesis regarding this issue (Sales department) with intervention mentioned too.

   You can now find and define a data or an information source, which you use to back up your thesis and anti-thesis and triangulate management reports (information sources).

4. Create a sentiment report for any organization or product of your choice. What decisions could be taken on basis of this? Also consider comparing this organization with a competitor.
5. Final question: How to enable the collective brain using concepts from the Hegelian inquiring system? What content are we talking about in the Hegelian inquiring system and what is the collective nature of the content, systems and services involved?

6. Pragmatism, social software and information services

6.1 Pragmatism

Singerian pragmatism is an epistemology and ethical theory stating that the value of knowledge should be expressed in terms of how the knowledge improves the human condition and, although people have to strive for truth, they will never reach the ultimate truth, and progress itself is regarded to be more important than truth itself. Malachowski (Malachowski, 2010) describes pragmatism as built on previous thoughts of philosophers C.S. Pierce, William James and John Dewey, with three commitments: 1) objectivity is historically situated, and none the worse for that, 2) knowledge has no foundations, and 3) philosophy needs to keep connected to first order inquiry, to real examples, and to real life experiences (pp. IX-X). Singerian pragmatist inquirers seek the creation of *exoteric knowledge*, or knowledge for "every man," *as opposed to scientific, esoteric knowledge* that as it matures, becomes relevant to an increasingly smaller audience. The artificial division of knowledge into disciplines and the reduction of complex problems into simple components inhibit the solution to social and management problems. Solving complex problems may require knowledge from *any* source and those knowledgeable in *any* discipline or profession. Attempts to achieve such kind of pragmatic exoteric knowledge have substantial organizational and human consequences. Courtney (2001; 28)(Courtney, 2001) says about this that Singerian organizations must bring in

> *"…multiple perspectives or worldviews and employing a holistic, systems approach in their thinking and decision-making processes. […]. A critical aspect of developing multiple perspectives is open, honest, effective dialogue among all relevant stakeholders in the problem involved. Managers in such an environment must be careful to respect the rights and viewpoints of the parties involved, and be open and honest to themselves in order to gain the trust of those who will be affected by the decision*". […] *In "real-life" situations, managing problems consists of at least three activities: (a) analyzing alternatives, (b) making decisions about which alternative to choose, and (c) successfully implementing the chosen alternative. The Lockean, Leibnizian, Kantian and Hegelian perspectives focus most strongly on (a) and least on (c); hence the "gap" so often deplored between analysis and action. Successful implementation depends first and foremost on the use of human resources and this means that we have to move from (a) to (c)."*

Obviously, multi-disciplinary work and the participation of many people, academic and practitioners is needed to improve mankind. Scientists are part of the game, but not the only participants, but society expects its scientists to work interdisciplinary and in close collaboration with problem owners, who have a clear stake in improving the human condition. Following this, one has to state how modern information

technologies can be of use here. For this, I see social software as particularly suitable. Social software, in contrast to business applications, is open for many, has information and use functions of everyone for everyone, and enables the collaboration of nearly everyone in problem solving. Social software does not only help in problem solving but also enables a new kind of practical knowledge named "the wisdom of the crowd". We discuss social software applications in section 6.2 and strategies for social listening in section 6.3

6.2 Social software

A review of literature on software for collaboration is given by Haefliger et al (Haefliger, Monteiro, Foray, & von Krogh, 2011). They say that social software may affect the interaction patterns between organizational members, create new opportunities for knowledge and information sharing, or unfold the disruptive and possibly change-inducing potential of so called "informational capabilities". Informational capabilities refer to an information technology's potential to alter the storage, transmission, and creation of information in an organization.

The term "social software" means software that supports group interaction. Today, social software, frequently annotated with Web or Enterprise 2.0, receives a lot of interest from managers due to its commercial use, increased network functionality, massive mobilization of users in some cases, and growing infrastructure capabilities, such as multi-media streaming online.

Social software affects the interaction between employees within, and individuals outside, the firm, such as members of user communities or customers. In many industries, users of technology, frequently organized in communities, are known to innovate independently of manufacturers (Von Hippel, 2007), and consumers have successfully contributed to innovation and product development organized by firms (Franke, Schreier, & Kaiser, 2010). Here, users and customers set up the governance structures for their communities independently of firms, often voice criticism toward firms and their products, or develop rival products in existing markets. Hence, in terms of strategic analysis, users and consumers can be "suppliers," "competitors," or "providers of substitutes".

There are three major implications resulting from the use of social software that favors a broader view of collaboration, extending beyond the company. First, users can assume several strategically important roles for the company beyond their obvious role as consumers of products and services. Second, social software shares with all information technology the capacity to change organizations in unpredictable ways, because it directly alters the way and the location where information is stored, shared, and created (Leonardi, 2007). The fact that most "outside members" of social software platforms are unknown to the firm makes it even harder to foresee how ICT will change the organization. Third, users rely on social software to organize within online communities that may or may not be supported by companies, and

"develop a life of their own". Understanding which interventions by the company will be perceived as beneficial or obtrusive is key to building lasting relationships with members of such platforms.
There are a number of recent contributions in strategy and organization theory that have addressed issues involving social software in the domains of strategy, technology, and community, see Table 6.1.

	View from inside the firm	**View from outside the firm**
Strategy	*Implement*	*Emergence*
Value creation	Inviting and empowering customers to contribute to product development	Strategic interaction with other users and learning benefits
Value appropriation	Firm's differentiated involvement in communities, dual licensing, selective revealing, better innovation performance	Availability and dissemination of assets under Open Source and Creative Commons licenses, or appropriation by user entrepreneurship
Technology	*Deploy*	*Self-expression*
Social software as a tool	Gaining access to creative users, utilizing their judgement and their know-how	Use of blogs and community participation for self-expression and identity building
Social software as a mediator	Platform-induced biases, groups and user-generated content and behaviour as "runtime effect".	Technology architecture signals value and suggests (self-)assignment of tasks and specialization
Community	*Harness*	*Belonging*
Leadership	Trade-offs between community founding and sponsorship, community leadership costly and complex	Central role of most achieved members of the community, social skills matter beyond technical savvy
Boundary	Cultural differences as challenge, risk of knowledge leakage	Firm involvement makes a difference in terms of contribution and motivation, yet firm recognition matters.

Table 6.1: A framework for social software. Source (Haefliger et al., 2011) with minor modifications

So what social software is specifically suitable for pragmatic uses, i.e. collaboration in problem solving and informing? We first describe a set of social software for only socialization and people networking after which we describe social software for entertainment goals. Both are not necessarily aiming at collaborative problem solving or knowledge creation, but they can be useful for finding appropriate people to team with. Finally we discuss some social software which aim at joint problem solving and knowledge sharing. For generating a quick and up-to-date list of social software, we checked Wikipedia, which has a great list of social software tools. The field, however, is advancing quickly and at the moment of publication this list is of course already obsolete. So please check Wikipedia regularly and contribute if you know of software and tools that are not available in their list. On October 18, 2012, we found the software we present below:

Social networking platform have specific location-based variants and more interactive event and action-oriented versions.

- General social networking platforms are: Facebook, Twitter, Pinterest, Instagram, Bebo, Chatter, Cyworld, Diaspora, Google+, Hi5, Hyves, IRC, LinkedIn, Mixi, MySpace, LAGbook, Netlog, Ning, Orkut, Plaxo, Tagged, Tuenti, XING, Yammer.
- Location-based social networks: Facebook places, Foursquare, Geoloqi, Google Latitude, Gowalla, The Hotlist, Yelp, Inc.
- Interactive and action oriented services: Engagement Advertising & Monetization: SocialVibe; Online Advocacy and Fundraising: Causes, Jumo, Kickstarter, IndieGoGo; Social Media Optimization: SocialFlow; Events: Eventful, The Hotlist, Facebook events, Upcoming,Yelp, Inc.; Social Media Gaming: Zynga, Empire Avenue.

We leave a discussion and analysis of this list to reader.

The October 2012 Wikipedia page gave us the following list of software and services that support *entertainment*.

- Game sharing: Zynga. Armor Games, Kongregate, Miniclip, Newgrounds
- Media and entertainment platforms: YouTube, MySpace, Cisco Eos, mtv.com, Qik, Vimeo, Dailymotion, Metacafe, Nico Nico Douga, Openfilm, sevenload, Viddler
- Virtual worlds:Second Life, Active Worlds, Forterra Systems, The Sims Online, World of Warcraft, RuneScape
- Livecasting: YouTube, Skype, Ustream, blip.tv, Justin.tv, Livestream, oovoo, OpenCU, Stickam

- Music and audio sharing: Pandora Radio, GrooveShark, Spotify, Guvera, Bandcamp, ccMixter, The Hype Machine, imeem, Last.fm, MySpace Music, ReverbNation.com, ShareTheMusic, Soundclick, SoundCloud, Turntable.fm, 8tracks.com
- Photography and art sharing: Instagram, Pinterest, Flickr, Picasa, deviantArt, Photobucket, SmugMug, Zooomr, Webshots

Key point for nearly all these plarforms is that they revised the classic broadcasting orientation (delivery of content one way from content owner to content consumer) to a two way interaction in which consumers can upload content for other and even easily deliver user-generated content, and the possibility of analyses of interest of users so that services can be fine-tuned and people with similar interests can easily find each other.

Next we have a large set of software and services for *collaborative purposes* from Wikipedia's social software page of October 18, 2012, and we added a few (Lucidchart, CMAP and ResearchGate):

- Brand monitoring, which also named social media measurement: Attensity, Statsit, Sysomos, Vocus, SocialFlow, Simplify360, Brandwatch, Webfluenz
- Blogs: WordPress, Blogger, BlogHer, Drupal, ExpressionEngine, LiveJournal, Open Diary, TypePad, Vox, Xanga
- Business reviews: Customer Lobby, Yelp, Inc.
- Community Q&A: ask.com, Askville, EHow, Quora, Stack Exchange, WikiAnswers, Yahoo! Answers
- Content Management Systems: WordPress, Blogspot, E107 (CMS), Drupal, Joomla, Plone
- Diagramming and Visual Collaboration: Creately, LucidChart, CMAP.
- Document Managing and Editing Tools: Docs.com, Dropbox.com, Google Docs, Syncplicity
- Information Aggregators: Netvibes, Twine (website)
- Microblogging: Dailybooth, FMyLife, Google Buzz, Identi.ca, Jaiku, Nasza-Klasa.pl, Plurk, Posterous, Qaiku, Tumblr
- Presentation sharing: Prezi, scribd, SlideShare
- Product reviews: epinions.com, MouthShut.com, Yelp.com, Cnet.com, Amazon.com product reviews
- Research/Academic Collaboration: Mendeley, Zotero, ResearchGate
- Social bookmarking (or social tagging): CiteULike, Delicious, Diigo, Google Reader, StumbleUpon, folkd, Zotero
- Social navigation: Trapster, Waze

- Social news: Digg, Stumble Upon, Chime.In (formerly Mixx), Newsvine, NowPublic, Reddit
- Wikis: PBworks, Wetpaint, Wikia, Wikidot, Wikimedia, Wikispaces, Wikinews

So that is a lot, and very likely not complete at the date I accessed Wikipedia (October 18, 2012) and very much not complete at the moment that you read this.

With wisdom of the crowd we mean taking into account the opinion, or wisdom, of a large groups of individuals (the crowd), rather than relying on a single expert. History has shown that the aggregated answers of a large group are often as good or better than the answer given by individuals within the group. This idea was first described in detail by James Surowiecki in his book titled "The Wisdom of the Crowds" (2004)(Surowiecki, 2005). While this process is not new, due to the rising popularity of the internet and Web 2.0 social media applications it has been receiving more and more attention. For example by websites as: Wikipedia, Ask.com and Yahoo Answers. However, when using the wisdom of the crowd, one has to be sure to take key elements into account.

- Firstly: diversity of opinion, every person should have his own opinion and the more diverse the "crowd" is, the better.
- Secondly: independence, the opinion of individuals in the crowd should not be determined by those around them.
- Thirdly: decentralization, one must not direct individuals in the crowd to work in a certain way, but rather let them work in their own way to find a solution.
- Finally: aggregation, there must exist a mechanism to aggregate all the individual judgments into a collective one.

Wisdom of the crowd is strongly related to the Singerian philosophy, as it seeks to find knowledge from the crowd hereby decreasing the power of a single expert on a topic.

There are also websites that exploit the "wisdom of the crowd" approach in a more commercialized way. These so called "prediction markets" are speculative markets which let people bet on the outcome of certain predictions. The current market prices of the stock determines the "probability" of the bet outcome to happen. Pennock, Lawrence & Nielsen (Pennock, Lawrence, Giles, & Nielsen, 2001) state that in many cases these predictions strongly correlate with actual outcomes. For example: the prediction market pays the holder of the stock $1.00 dollar if the event ("Obama wins the re-elections") happens and $0.00 dollar in case the event does not happen. Therefore if the stock price is $0.80 dollar this would mean that the crowd thinks there is an 80 percent chance that President Obama will be re-elected. A practical and useful prediction market service is www.ipredict.co.nz.

6.3 Social listening

Given the importance of the views and opinions of the crowd and the huge number of these on the Internet, organizations may profit a lot form developing capabilities of efficiently listening to the crowd. This section discusses how this can be done efficiently.

6.3.1 Social media vision and objectives

There are currently over 500 million users on Twitter and over 1.11 billion on Facebook and these numbers are growing. Not engaging in social media could result in missed opportunities. But, being successful as an organization at social media is not easy. De Vera and Murray of Greenlining.org state in their very useful guide to social media listening that:

> *"We'd love to be able to tell you that Greenlining's experience with social media has been an unbroken string of successes, but to be honest, it has been a bit of a roller coaster ride. We started out in 2007 by hopping on the bandwagon without a clear vision in mind. We created social media accounts and posted items here and there, but made the critical mistake of not understanding why we were using these tools. Instead of viewing social media as necessary tools in our communications strategy, we used them sporadically and randomly, and at times neglected them. We mistakenly thought we would instantly reap the benefits of social media after posting a few things, a common mistake in the nonprofit sector. (….) In 2011, after years of frustration, we approached social media with a purpose. Being intentional meant three main things to us: (1) having goals and a strategy (2) investing resources (staff and money), and (3) believing in social media as a valuable tool to achieve organizational goals. (….). The result was dramatic. In less than one year, we grew significantly, tripling our Twitter followers, for example, from a humble 500 to 1,500 in 10 months. But beyond just numbers, we were accomplishing something far greater: We were listening to what people were saying about our issues and building relationships as a result"* (Source: "The art of listening" from www.greenlining.org).

Developing a well thought-out vision is the first step towards an effective social media strategy. Before you start using any social media tool, you have to return to your own organization's goals, and ask the following questions when developing a vision:

- Who is your target audience?
- What do you specifically want to accomplish by using social media? This can be for instance raising awareness about your organization, better brand recognition, developing meaningful relationships with online influencers, getting people to take action, fundraising, increasing attendance for an event, and educating and informing audiences.
- How do social media connect with your overall organizational vision and mission?

- How much time and resources will you invest to manage social media activity?

As a rule of thumb, have three to five goals for a vision statement, which answer the questions: "Why social media?" Such a vision statement may be phrased as "Through the use of social media, our organization seeks to…." for example…

1. Engage with the general public to raise awareness about issues
2. Create conversations about our work to challenge the mainstream discourse
3. Listen to our audiences to gain their trust and insight
4. Build and develop relationships and brand our organization as a leader in the field

This vision has to be transformed in concrete SMART objectives. These objectives will allow you to constantly evaluate your strategy's effectiveness and make changes as needed. SMART objectives are…

- Specific: Answers the question, "What exactly are we going to do and for whom?" This identifies what will change in concrete terms, by identifying the particular population and/or setting (tools), and the specific actions that come as a result.
- Measurable: Objectives should be quantifiable and easy to measure as a means for evaluation.
- Attainable: Objectives should be feasible within a proposed time frame with the resources and support available.
- Relevant: Your objective, strategy, and overall organizational vision and mission should align.
- Timely: Good objectives have deadlines. Ask: When will this be accomplished?

The following template provides an efficient way to conceptualize SMART objectives.

"BY <When, date>, we <Who> will <What, attainable, relevant>, from <current measure, number rate, percentage> to <future objective number, rate, percentage> through <HOW>".

An example of SMART Objectives can be:

- By the end of July, maintain consistent activity on Twitter by increasing our daily tweets from 1-2 per day to 5-6 per day. Daily tweets will consist of original content, replies to other users, and retweets of relevant content.

Key performance indicators (KPIs) help you track your progress. Metrics are the numbers that you capture. There are different KPIs you can measure, such as: total amount of followers and fans, your growth rate over a period of time and the reach you have on Facebook and Twitter. When you measure, look into specific details, such as the types of posts that were the most effective, and take note of them. See Tables 6.2 and 6.3 for examples of KPIs and metrics.

Goal	Objective	Strategy	Tactic	KPI	Metric goal
Engage our clients to discuss over services	Regular comments on Facebook posts and Twitter mentions	Pose questions on Facebook. Participate in chats on Twitter to mention other users.	Ask questions on current issues (like new menus and ideas for activities). Listen to the hashtags on Twitter	Number of Facebook comments, number of Twitter mentions	50 Facebooks comments, 150 Twitter mentions per week

Table 6.2: KPIs for a social media strategy

Goal	Metric	Jan-12	Feb-12	Mar-12	Apr-12	May-12	Jun-12	Jul-12	Aug-12
Active community engagement	Number of comments on posts	7	7	12	5	16	12	9	12
	Average number engaged users per post	14	9	9	13	10	22	16	32
	Number of likes on posts	89	70	134	89	111	129	159	296

Table 6.3: Example: Facebook Metrics

Keep all of your metrics in a central spreadsheet. An Excel document will help you create graphs and tables of your social media progress throughout a given time period. Keep a folder of all social media analytics reports and dedicate time to analyze and understand the numbers to develop better social media tactics and strategies for the future.

6.3.2. How to "Listen"

For listening well, take time to develop a basic understanding of the issue(s) that your organization works on. Get answers to the following questions:

- What is the vision and mission of the organization?
- What is the purpose of the different programs and initiatives?
- What are the past and present goals of these programs and initiatives?
- What are the current issues that the programs and initiatives are working on?
- Who are the most influential individuals, organizations, and thought leaders on these issues?

Each of your program and issue areas should outline their purpose and goals, and have a list of key concepts (buzz terms) and influencers. Having a list of buzz terms and influencers that will aid you in determining what people are talking about online. Concepts allow you to identify trending topics on Twitter, and influencers will provide you with a group of key individuals and organizations to follow on Twitter and like on Facebook. See Table 6.4 for examples of Buzz terms for Grand Café.

Issue area	Purpose	Goals	Current issues	Buzz terms
Musical events program	To develop idea for a musical event each Friday with high potential but less known musicians and bands	The goal is to increase Grand Gafés as a place for people searching for new musical experiences	Grand Cafés image is too much foot and drink only and less attractive for younger people	Music styles, new bands, different music, young music, innovative music, youth culture,

Table 6.4: Issue areas and buzz terms for topics and influencers detection for Grand Café

Follow a manageable number of accounts and other popular media, like CNN, BBC news and newspapers. Listening to traditional media sources will provide you the opportunity to immediately share important news related to your organization's work and contribute to mainstream conversations. Useful tools for applying your buzz terms on are:

- Twitter Search – built right into the Twitter platform, this tool lets you search what keywords are being discussed in real-time on Twitter through the use of hashtags
- Netvibes – a tool that allows you to input a keyword and see if it is being discussed on Twitter and other various social media such as blogs
- Facebook Graph Search – the new search tool built into the Facebook platform that allows you to search user accounts and hashtags to listen to what is being said about particular topics

After searching, aim to get around 4-5 buzz terms for each of your programs/issues that you can closely monitor. Often, you will find hashtags related to these buzz terms. Develop a set of criteria that outlines what is considered a trending topic for your organization to help you prioritize hashtags to follow closely. Keep in mind that the buzz always changes. With the 24-hour news cycle, new issues will take the forefront of mainstream media every day, so set a schedule to regularly update your terms.

6.3.3. Communicate on Social Media

Make your content understandable to a wide audience. Invite people to join the conversation — don't leave them out. On Twitter, you have a 140-character limit per tweet, so make the most of every character. The use of well-known acronyms is acceptable and encouraged, such as "CA" for California or "POTUS" for President of the United States. If you want to link your audience to important websites or news articles, use a link shortening tool like bit.ly or tinyurl to shorten website addresses to 20 characters and track click-through rates.

Figure 6.1: Anatomy of a (re)Tweet. A retweet by Joe Peppard of a tweet by John Torp with a "mention" of another organization and person (via the @) and a "hashtag" (#) of twitter topic Forbes with a "link" by a "tinyurl" to another site.

Constant activity on social media establishes your organization as a legitimate source of information on issues your organization works on. Consistency gains trust from other social media users. Inconsistency, on the other hand, not only highlights an organization's lack of presence and recognition, but might also cause users to forget about you altogether. To avoid being forgotten, set goals for daily activity. Realistically, this is dependent on your staff capacity, but successful social media accounts are consistently posting and engaging. Table 6.5 gives a media activity plan.

Daily	Weekly	Monthly
Photo's, video, quote of the day. Articles and stories about musicians. Statistics of the day.	Song of the week. Weekly ask to increase number of followers. Announcements of event.	Overview of the month. Review of what career progress each past performing artist or band did. Idea for the future program.

Table 6.5: Social media activity plan for Grand Café

Once you create a flexible schedule that works for your organization, test it out and observe what types of content work well on particular days and times. For example, Mondays may not be great days to roll out online campaign content since a lot of people are easing themselves into the work week. As a result, you might want to focus on promoting this content during the middle of the week. The prime hours to post on weekdays are between 11a.m. – 2p.m. EST. For an example see Table 6.6. If you are already using Facebook and Twitter, a content management tool will be helpful to manage both at the same time. There are a variety of free and paid tools available for social media users, such as Hootsuite, Seesmic, and Tweetdeck. If you are starting out, try Hootsuite first. The basic version of Hootsuite is free and allows you to manage your Facebook and Twitter accounts at the same time. You can create dashboards of all of the news feeds and streams you find relevant. Special built-in features allow you to focus your streams on specific keywords, hashtags or lists of users. You also have the ability to schedule social media posts throughout the day, week, and month, giving you the flexibility to ensure you have regular activity on Facebook and Twitter. Hootsuite also provides basic analytics that help you measure your social media impact.

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Evaluations of the week. New menus and activities	Announcement of Friday event	Announcement of number of attendants and space reservations	Last preparation information for event	Online twittering wall during event	Evaluation of Friday event	Nothing

Table 6.6: Posting schedule for Grand Café

6.3.4. Build an Audience and Following

The quality of your followers is more important than the quantity. Start building your audience with your immediate networks. Publicize your Facebook and Twitter accounts amongst your staff: Send out an email instructing them to like your Facebook page and Twitter account, then have them go a step further by asking them to share the accounts with their own networks. Tap into the close relationships and partnerships your organization has with other groups and individuals, and create a simple ask to have them follow you and publicize your accounts with their own networks as well. After tapping into your close relationships, your audience should organically build and expand.

Continue building your base by following and identifying organizations and influencers that do similar work. A compiled list of partner organizations and influencers will come in handy here. Search for these organizations and individuals on Facebook and Twitter, and follow/like them if they have accounts. Once you follow these accounts, you can check out who they follow and like, thus expanding your list of organizations and individuals to follow. Next, create criteria that will help you determine whether you should follow certain accounts. You want to follow accounts that are active and engaging.

On Facebook and Twitter, listen to conversations by searching hashtags. If there is already a conversation on a certain topic (i.e. Affordable Care Act, #ACA), then you can use Twitter Search and Facebook Graph Search to find all of the posts related to the conversation. This allows you to find users to reply to and engage with in conversation. This is an organic method of building online relationships with organizations and individuals that you may have never interacted with before.

Brand your hashtag with something easy to remember and relevant to what your event or campaign is about. For example, Greenlining hosts an annual Economic Summit and their theme for 2013 was "Vision 20/20: Our Legacy, Our Horizon." The hashtag they chose for the event was "#vision2020." They ran a campaign in 2012 to pass a bill in the California legislature, and used the hashtag "#SB1233," the number of the bill. Choose unique hashtags to make sure that the conversations are focused.

Since hashtags are public, many may already be in use, which can be a good thing. As long as the conversation is related to what you're talking about, then it's completely okay to use the existing hashtag. In fact, you'll be offering new perspectives to ongoing conversations. However, if existing conversations are out of scope, then you may want to find another unique hashtag to brand your event or campaign.

The number of accounts you like and follow may become overwhelming to the point that your newsfeeds are overloaded with information. For example, what if you're searching through your news feed for content related to a specific issue area? It can be frustrating having to sift through the numerous updates on your Twitter and Facebook newsfeed to find relevant content. Creating Twitter lists can make listening easy. For example, you can maintain lists for each of your issue areas (e.g. health, education, economic justice), lists for specific kinds of people (e.g. reporters, elected officials, staff members, etc.), and lists

for different media sources (e.g. national, state, local, ethnic). By creating lists, you can sort user accounts into different groups and create focused streams of posts. Additionally, creating Twitter lists allows other users to be notified that they have been added to a list, which gives them an incentive to give you a mention or follow you back. Users can also subscribe to your lists which can signify user appreciation for your social media activity.

6.3.5. Develop a Social Media Policy

When the social media fever spreads to your staff, this excitement also poses some challenges and potential dangers. For example, how do you manage the content that various staff members post and send out? What is considered an appropriate post as a nonprofit employee? Where do the lines between the personal and professional lie? Developing a social media policy allows you to clear up any confusion by giving your organization guidelines for what should and shouldn't be done on social media. A good policy outlines your organization's values and social media goals, and indicates the roles and responsibilities that staff members have. It is good practice to have at least one person focused on developing and implementing your social media strategy. In addition to defining roles, consider creating a system of checks and balances. Do your posts have to get approved by a superior? If so, what kinds of posts? Who speaks officially about your organization, and about which topics? One way to define expectations is to create a "Do's and Don'ts" of social media for your organization like in Table 6.7.

Do	Don'ts
1. Be professional in and out of work. What you share still reflects on the organization.	1. Post anything that will jeopardize the organization
2. Be responsible. Remember that what you publish will be attributed to you	2. Post what you don't want the public to see; if you don't want your parents to see it, don't post it!
3. Be transparent. Use a disclaimer that states views are your own.	3. Post anything obscene, offensive, or profane (without very good cause)
4. Be respectful and civil	4. Reveal confidential information, especially if it is work-related
5. Think before you post. Keep in mind the global audience you're sharing with	5. Intentionally mislead or misrepresent the truth
6. Make sure you can back up what you say	6. State wild ideas as truth

Table 6.7: Some organizational social media policy do's and don'ts.

6.4 Assignment

1. Following Table 6.1, multiple perspectives and issues can be taken for any social software use by an organization. Specify these issues following the structure of Table 6.1.
2. What social software can Grand Café use for what tasks to improve internal and external collaboration?
3. Go to prediction market http://www.ipredict.co.nz/ and try out what Grand Café could use ipredict effectively for. What are the opportunities and risks of using prediction markets for management decisions?
4. Develop a social listening strategy for any organization of your choice? How can socialmention.com be used for this purpose? For this assignment be sure that your approach carefully matches with the organizational vision and strategy.
5. Final question: What is the content of the collective brain following a pragmatist inquiring system? To what extent is the content collective? To what extent do you think that social software may be less collective as it assumes.

6.5 Further reading

On social software and the firm I recommend Haefliger (Haefliger et al., 2011) which is an introduction to a special issues on social software and organizational strategy in Long Range Planning. On pragmatism I recommend Churchman's classical book (Churchman, 1971) or Malachowski (Malachowski, 2010), which are both fundamental and philosophical. A more practical story is given by Courtney (Courtney, 2001). For the section on social listening (6.3), De Vera and Murray's very useful text from www.greenlining.org is recommended. This text also includes many references to useful tools and practices.

7. Organizing the collective brain

7.1 Organization and information

After having given conceptualizations of information and their practical implications in the previous chapters, we now discuss information management in its organizational context. To organize this discussion, we chose the IT-alignment model (Pennock et al., 2001). This model focuses on the relevance and contributions of IT investments for organizational strategic and operational requirements. Most organizations can only keep existing when their strategic policy, the IT-policy and the IT-systems are well aligned. In some cases the IT and systems policies do not correspond with the strategic policy or organizational processes. In those cases the organization does not work efficiently and effectively. To solve this problem, the coherence of IT resources and the organization in general has to be analyzed and second the internal structure of the IT resources (called the IT architecture) has to be well organized. For this we developed the organizational information management model of Figure 7.1.

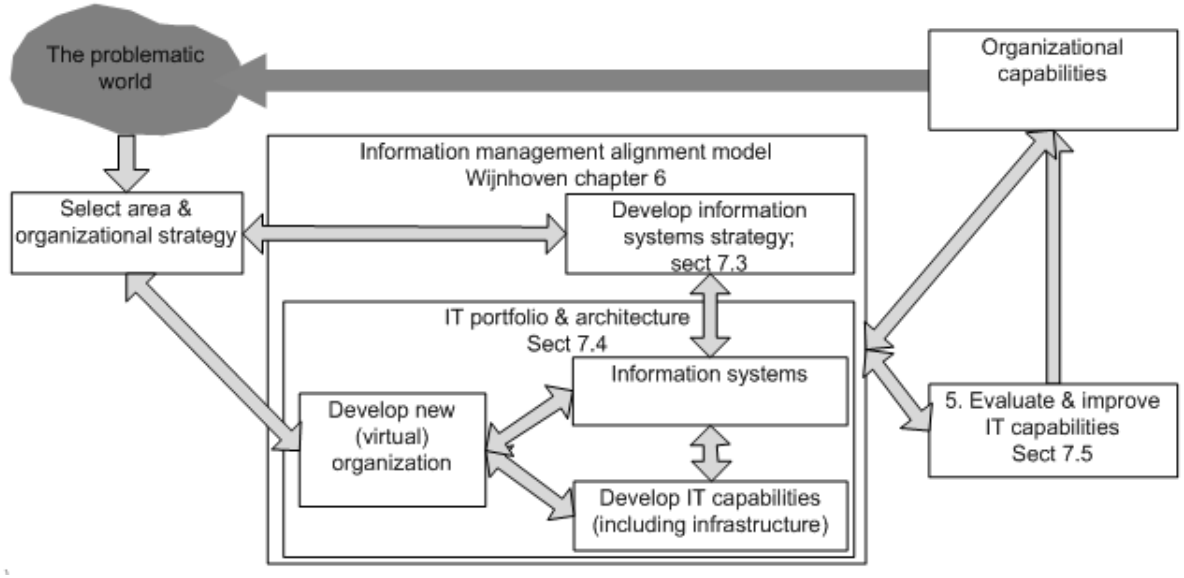


Figure 7.1: The extended alignment model and sections of this chapter.

7.2 Goal of this chapter

The goal of this chapter is to introduce problems of developing an IT strategy for an organization. For this, section 7.3 introduces key concepts for developing an IT strategy aligned with a company's needs. An IT strategy gives the guidelines for the development and adoption of IT means, which together are

also named the IT portfolio. An IT portfolio is split in existing IT services (like databases, applications and security services) and the IT project portfolio (De Haes, Gemke, Thorp, & Van Grembergen, 2011). We discuss the idea of IT alignment as a policy principle in section 7.3. Next we discuss the IT service portfolio in section 7.4 and we discuss the IT project portfolio in in section 7.5. Section 7.6 gives multiple ways of developing an aligned strategy. Section 7.7 gives assignments and 7.8 gives suggestions for further reading. The development of these IT policies is comparable to the purposeful development of a collective brain, i.e., not just being there as some material but as the enabler of organized thinking about the relevant world. This organized thinking does not only have people with their brains but also includes IT tools that extend people's capabilities by storage and memory, reasoning, collaboration, negotiation, problem solving and decision making.

7.3 IT strategy and IT-organization alignment

Bocij et al (Bocij, Greasley, & Hickie, 2009) state that a good alignment of IT policies is difficult to achieve because:

1. Organizational strategies are often difficult to identify and there may be a difference between the existing strategy and the intended strategy.
2. Organizational strategies are never stable and have to constantly adapt to changing environmental circumstances.
3. The resulting IT means have a need for change on their own, given developments on the IT market place.
4. It is often hard to make an IT strategy that satisfies the needs of all organizational stakeholders, and consequently IT portfolios are not always consistent with the IT strategy.

For these issues, figure 7.2 structures the alignment problems.

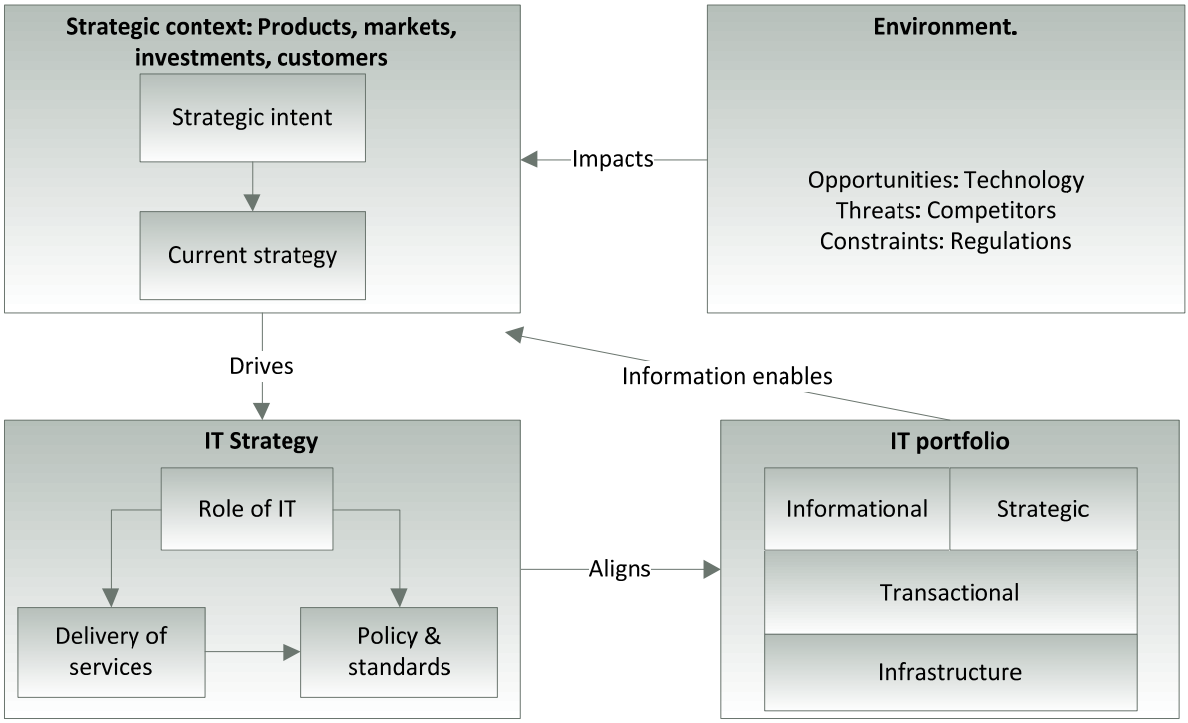


Figure 7.2: Business and IT alignment. Source (Bocij et al., 2009) with some modifications

As this figure 7.2 explains, a company's strategy must be in line with environmental needs but cannot be designed fully without an understanding of the company's current strategy and resources. The IT strategy should be derived from this business strategy, and includes statements regarding the role of IT in the strategy, the types of IT services that the company expects, and needs from IT and possible policies and standards to realize a coherent, synergetic and efficient IT. This may result in a portfolio of IT architecture components and IT projects (De Haes et al., 2011). Some of these portfolio components focus on information content (like its quality, integration and availability) and others may focus on new IT strategic issues and policies, others with improved business processes or transaction services, and finally some may focus on IT infrastructure elements to support applications and processes (like security and networks). As one can see, therefore, IT is not only derived from business and strategic needs, but it also enables new business strategies.

Regarding the role of IT in the firm, Ward and Peppard identify a strategic grid of information systems application. See Figure 7.3.

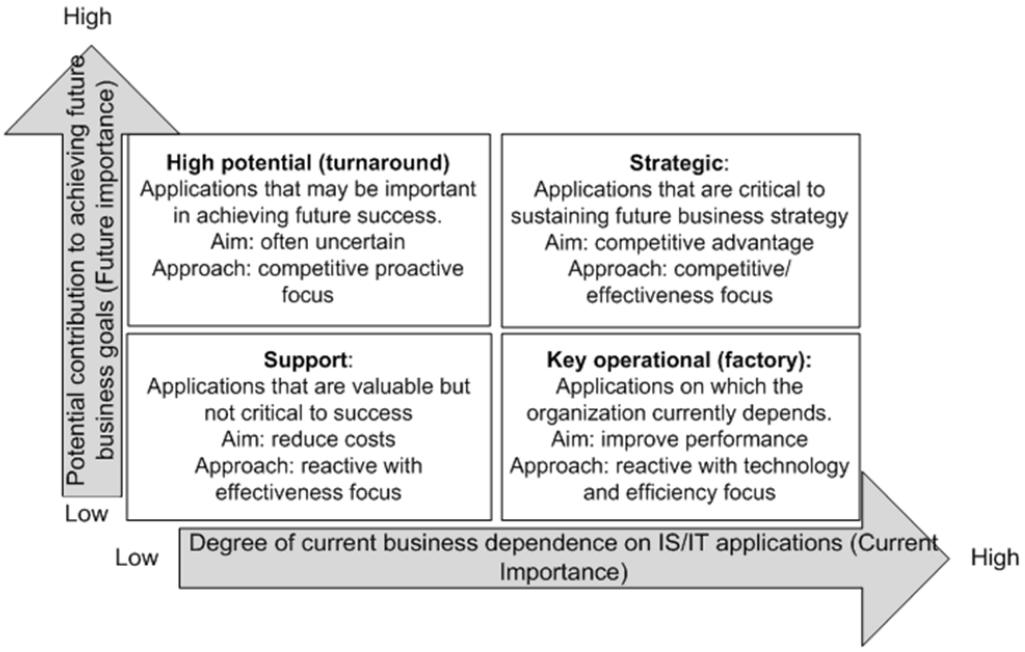


Figure 7.3. Ward and Peppard's strategic grid. Source (Bocij et al., 2009) with modifications

Each system or application can be assessed in strategic or operational terms. These assessments drive IT investment decisions and thus also priorities among possible IT projects.

7.4 IT Architectures

For organizing the set of IT services that an organization needs, these services should be well linked to specific business needs and be well supported by software, technology and people. Such an organization can be described in an architecture language which is an organization chart for the IT resources of an organization. ArchiMate is one of the leading languages to represent such an organization. It includes three layers for the IT architecture: the business layer, the application layer and the technical layer.

- The business layer provides products and services to consumers and clients. In business administration terms we call this the front-office. This layer consists of, for example, a number of products and services. This layer also has a number of supporting services. For example: with the production of bicycles you could have a sales process and a supporting process, like an HRM policy to ensure the availability of the best people.
- The application layer supports the business layer with application services which are realized through software. As an application you could have an HRM service with as supporting system an HRM application.

- The infrastructure layer consists of all the infrastructural services which are needed to run these applications. For example: the HRM database needs physical storage capacity. You also need communication which makes it easier to add new data to the database, security, privacy and networking.

See Figure 7.4 for the ArchiMate layers.

- The "service" is a unit of functionality that some entity (system, organization, department) makes available to its environment. For example a "health check" is a service offered by a physician.
- When combined with layers, we create a layered view of the enterprise.
- Layers make services available, which are used by others layers in the enterprise architecture.

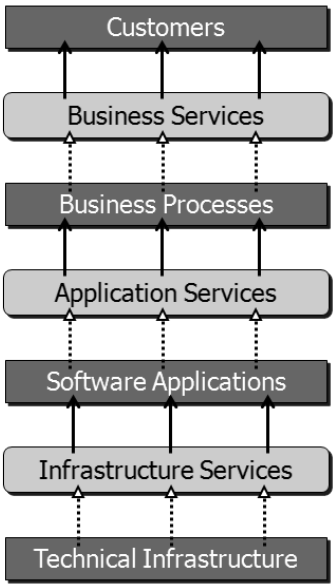


Figure 7.4: ArchiMate layers

Next to the horizontal layers, ArchiMate distinguishes vertical aspects. These aspects are: the information aspect, the behavior aspect and the structure aspect.

- The information aspect takes care of the information objects and the data structures. This is also called the "passive structure".
- The behavior aspect organizes the actions that transform information to meet the information need.
- The structure aspect contains an organized group of persons, applications and information technology which perform the behavior that is necessary to fulfill the information need.

When you follow the above mentioned structure you will have a matrix with 9 cells, because you have 3 horizontal layers and 3 vertical aspects; see Figure 7.5. A list of representation objects and their location in a 3*3 ArchiMate matrix is given in Figure 7.5.

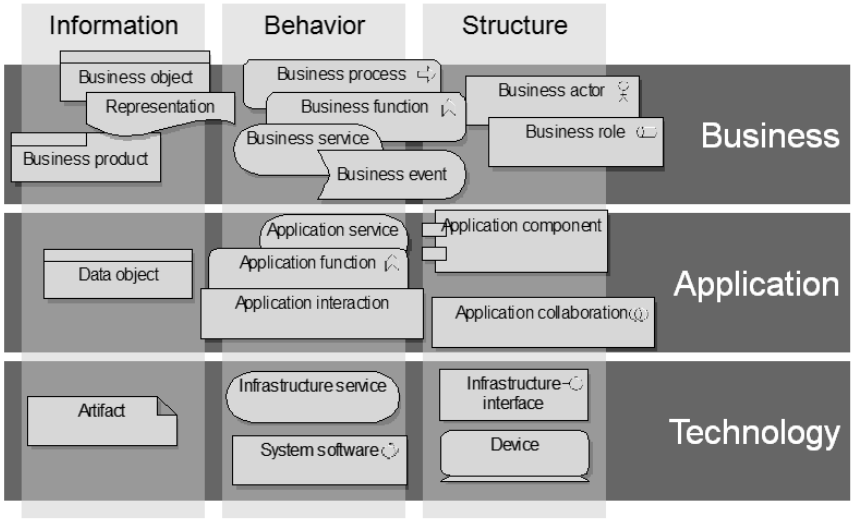


Figure 7.5: Objects in the ArchiMate framework

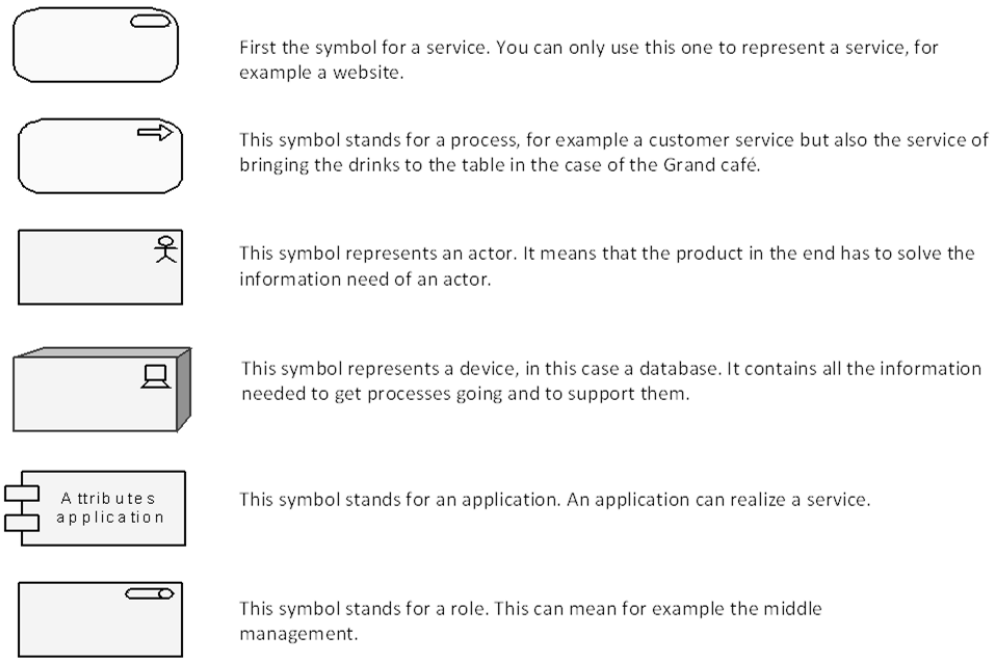


Figure 7.6 Some ArchiMate representation techniques

A key idea behind ArchiMate is that the aspect and layers component have causal relations to each other, and thus impact each other. The “used-by” and “realization” relations are represented by different causal

arrows. ArchiMate has two other (causality) representations, triggers and information flows. Additionally one can represent relations between architecture components also in terms of parts and whole. See Figure 7.7 for an overview of ArchiMate relations.

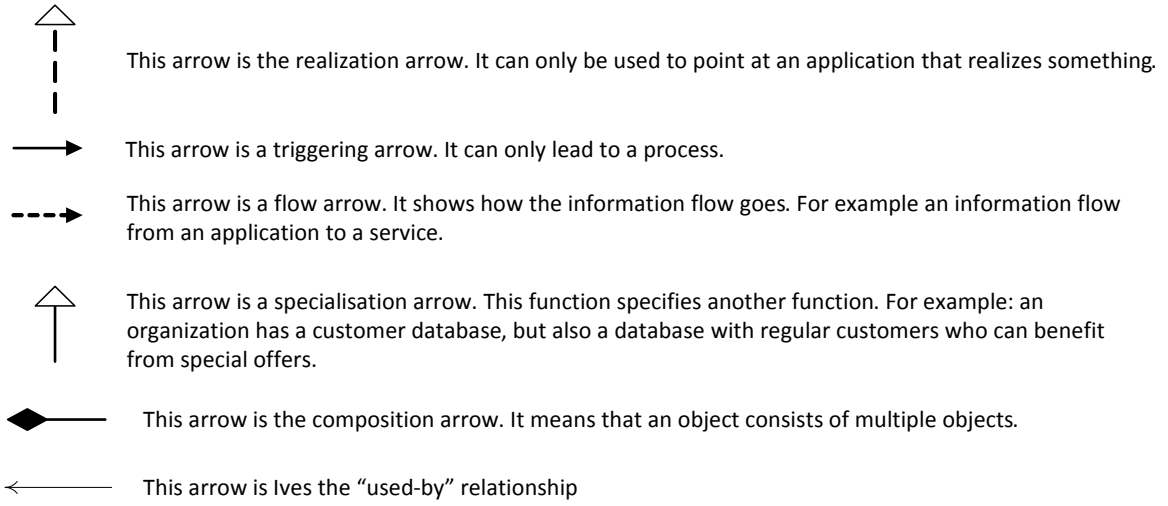


Figure 7.7: ArchiMate relations

For the practical identification of required applications, processes and infrastructure components for Grand Cafe, the following chart shows already that a lot has to be checked and planned to make an effective information management organization for Grand Cafe. Because of this complexity, we leave the differences between several layers out of the diagram. We also do not strictly follow the ordering of the aspect (passive, behavior and structure), but we do represent an enterprise architecture as an organized system of resources at Grand Café. See Figure 7.8.

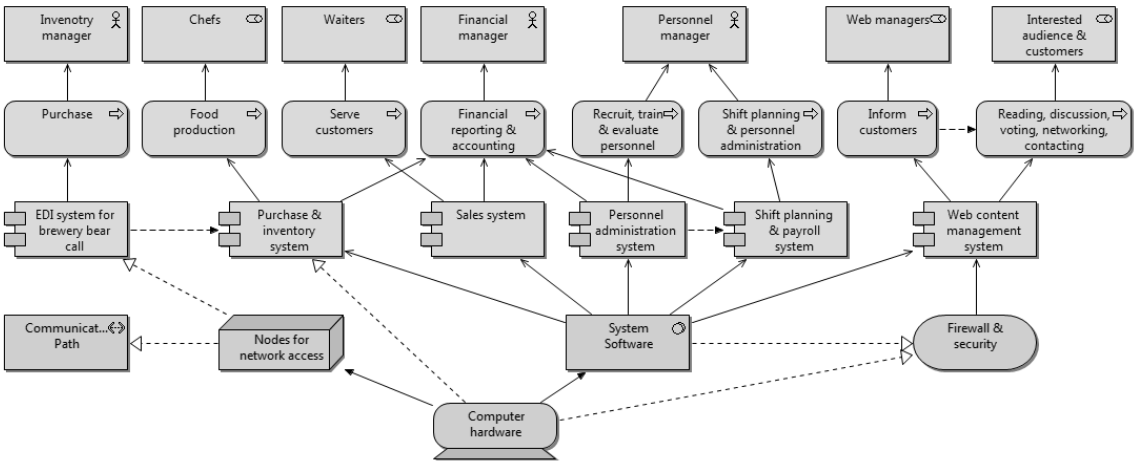


Figure 7.8: ArchiMate model for Grand Cafe

As you see from 7.8, many business processes are taken by different actors, who need several applications to let them run well. The beer-ordering application is owned and managed by the brewer, who is also responsible for its security, and thus Grand Cafe only needs to set up the network access to the brewer to realize efficient ordering processes. The transaction information from the ordering system is send over to the purchase and inventory system, which is used by the financial manager to check the bills and integrate the data in the financial reporting. Grand Cafe allows external to view information that Grand Cafe publishes on its website and they also allow their audience to discuss and interact virtually. This requires setting up a firewall and security service to be realized by several kinds of system software.

7.5 IT project portfolio management

For assessing the effectiveness of the IT portfolio, Buschle and Quartel (Buschle & Quartel, 2011) suggests to use the ArchiMate framework and analyze the contributions of IT components for business objectives. For this purpose they extended ArchiMate by value constructs. Figure 7.9 gives an example of such a value-extension of ArchiMate of an electricity power supplier who has three business services (one for the optimization of its own network, and two for advising clients). The first service aims at capacity optimization (too much is too- expensive and too less means lost business opportunities). The other two aim at increasing the client satisfaction and customer loyalty. Serving these objectives well is expected to increase profitability in the longer run.

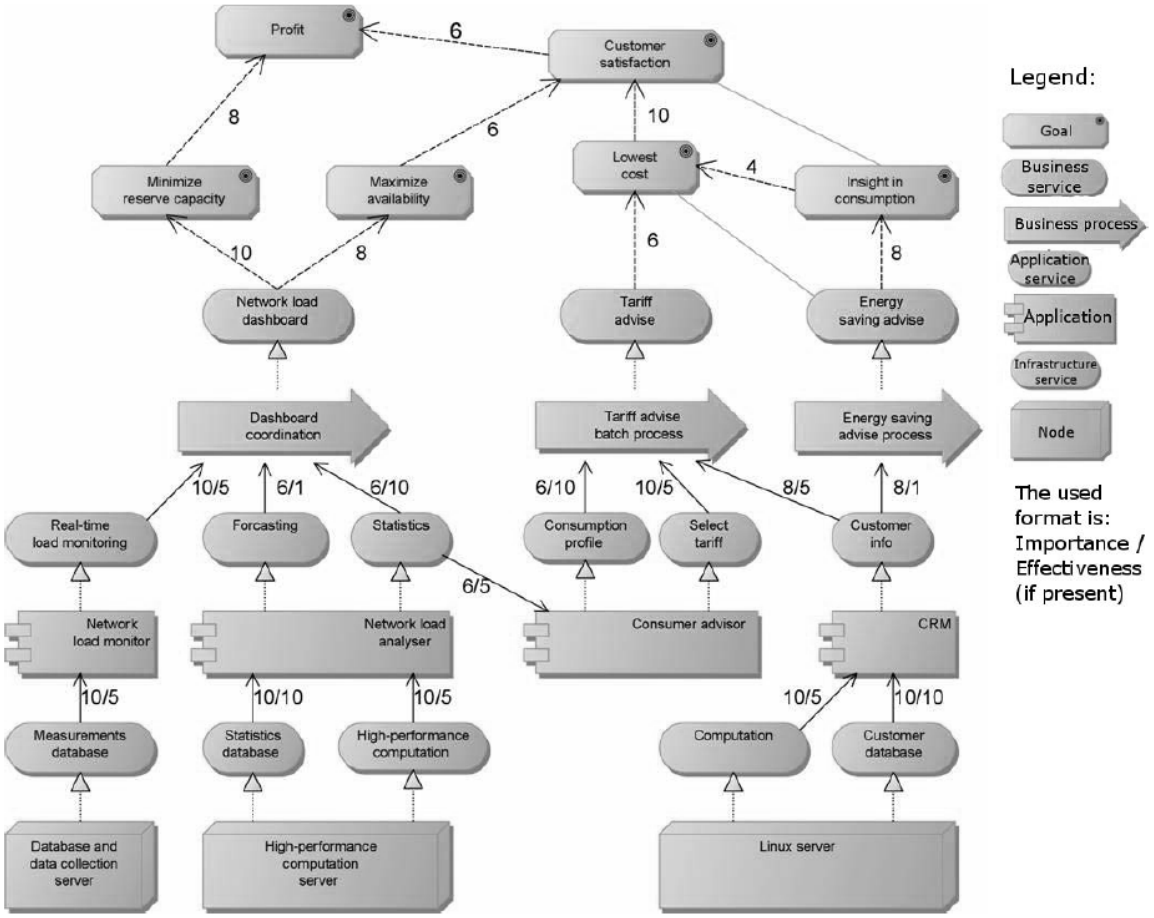


Figure 7.9: Example of Buschle and Quartel's ArchiMate valuation of IT components in an organization. Source: (Buschle & Quartel, 2011)

Figure 7.9 expresses causal relations in ArchiMate by the *importance* and *effectiveness* of a component for another component in the Architecture by a 1-10 scale. Thus Customer satisfaction is rated as 6 on the importance scale for Profit. The statistics database is of highest importance and effectiveness for the Network load analyzer. On basis of this they can produce project priority plans as in Figures 7.10 and 7.11.

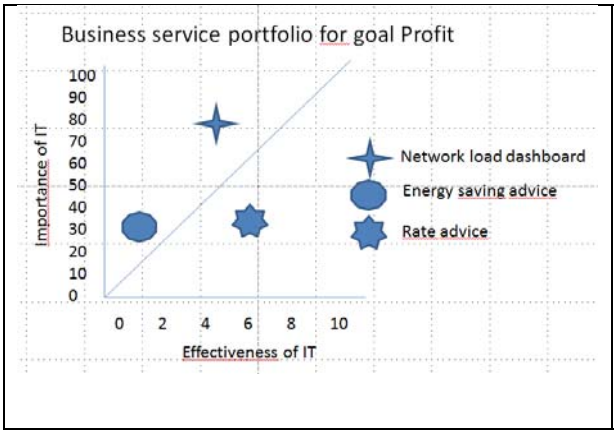


Figure 7.10: Profit portfolio corresponding with Figure 7.9. Source: Buschle & Quartel, 2011 with modifications

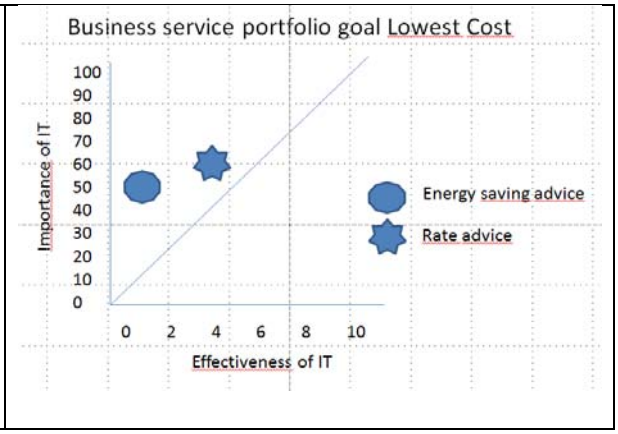


Figure 7.11: Lowest cost portfolio corresponding with Figure 7.9. Source: Buschle & Quartel, 2011 with modifications.

Of course, more metrics can be used to evaluate IT. A well-known and popular classification of metrics is based on the so-called balance score card, which gives four main value areas for a company, which need concrete metrics in an assessment (Kaplan & Norton, 1992). These value areas need some kind of balance, and an over-emphasis on only one dimension may bring the company at risk.

7.6 Further study

Lankhorst et al (Lankhorst, 2009) gives an good overview of the ideas and applications of ArchiMate. A clear practical tool and explanation is given by archi.citis.ac.uk. In ArchiMate one can use the stencils of Visio, which can be downloaded at www.opengroup.org/archimate/doc/ts_archimate/apdxa.html. Alternatively, you can use the Archi tool from archi.cetis.ac.uk, which can be downloaded for free.

A good example of working with the balanced score card in IT policy development give Van der Zee and De Jong (Van Der Zee & De Jong, 1999). Bocij et al (Bocij et al., 2009) and Laudon and Laudon (Laudon & Laudon, 2012) have useful introductory books that describe IT in organizational contexts. For strategic information policies I recommend Ward and Peppard (Ward & Peppard, 2002). The ArchiMate-based IT project portfolio method is well explained by Buschle and Quartel in (Buschle & Quartel, 2011).

7.7 Exercises

1. SmileYou is growing fast, but as an organization it has not much coherence yet and may fall apart soon. Consequently Mary-Ann wants to develop an organizational strategy that aims at tighter coherence of all shops and warehouses in all countries. In each country, one shop and one warehouse will be set up as the intermediary (the importer) between SmileYou headquarters in Amsterdam and

the local shops. Given this a new enterprise architecture diagram and specification of the systems for each shop, warehouse, importer and headquarter are needed. Make such a structure following the alignment model and next create the related ArchiMate diagram.

2. At a university a large set of different information systems exist, like the e-learning system, the student registration system, the grading system, the course planning and room planning system, and the budget and payment system. Create a relevant ArchiMate diagram for this situation that also aims at maximum reuse of the different systems resources, data and functionality.
3. For Grand Café the following questions can be raised and answered by you:
    a. How would its strategy and strategic intent fit with its IT strategy and IT portfolio?
    b. How can Grand Café integrate its portfolio into an IT architecture?
    c. Analyze the contribution of ArchiMate components to Grand Café's objectives and generate project and investment priorities.
    d. Develop an improvement plan consisting of a set of projects with relations, priorities and budgets given a limited development budget of 2 million euros for the next 3 years.
4. Final question: what is the content and collective nature of the collective brain in an organizational context?

References

Anderson, F. D., & Prelli, L. J. (2001). Pentadic cartography: Mapping the universe of discourse. *Quarterly Journal of Speech, 87*(1), 73-95.

Barnard, C. I. (1968). *The functions of the executive* (Vol. 11): Harvard Univ Pr.

Berger, P. L., & Luckmann, T. (1967). *The social construction of reality*: Doubleday New York.

Bocij, P., Greasley, A., & Hickie, S. (2009). *Business information systems: Technology, development and management*. Harlow, UK: Financial Time/Prentice Hall.

Buschle, M., & Quartel, D. (2011). *Extending the method of Bedell for Enterprise Architecture Valuation.* Paper presented at the 15th ieee International Enterprise Distrbuted Object Computing Conference.

Churchman, C. W. (1971). *The design of inquiring systems: basic concepts of systems and organization*. New York: Basic Books.

Courtney, J. F. (2001). Decision making and knowledge management in inquiring organizations: toward a new decision-making paradigm for DSS. *Decision Support Systems, 31*(1), 17-38.

De Haes, S., Gemke, D., Thorp, J., & Van Grembergen, W. (2011). KLM's Enterprise Governance of IT Journey: From Managing IT Costs to Managing Business Value. *MIS Quarterly Executive, 10*(3), 109-120.

Denzin, N. (2009). *The research act: A theoretical introduction to sociological methods*: Aldine De Gruyter.

Evans, J. R. (2013). *Business analytics: Methods, models, and decisions*: Pearson.

Fischer, P., Lea, S., Kastenmüller, A., Greitemeyer, T., Fischer, J., & Frey, D. (2011). The process of selective exposure: Why confirmatory information search weakens over time. *Organizational behavior and human decision processes, 114*(1), 37-48.

Franke, N., Schreier, M., & Kaiser, U. (2010). The "I designed it myself" effect in mass customization. *Management Science, 56*(1), 125-140.

Haefliger, S., Monteiro, E., Foray, D., & von Krogh, G. (2011). Introduction to Social Software and Strategy. *Long Range Planning, 44*, 297-316. doi: http://dx.doi.org/10.1016/j.lrp.2011.08.001

Halpin, T. A., & Nijssen, G. (1995). *Conceptual schema and relational database design*: Prentice-Hall, Sidney.

Hartnack, J., & Hartshorne, M. H. (1967). *Kant's theory of knowledge* (Vol. 61). London: Harcourt, Brace & World.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly, 28*(1), 75-105.

Hoffer, J. A., Prescott, M. B., & McFadden, F. R. (2007). *Modern database management*: Pearson/Prentice Hall.

Hofstede, G. (1981). Management Control of Public and Not-for-Profit Activities. *Accounting Organizations and Society, 6*(3), 193-211.

Huenemann, C. (2008). *Understanding rationalism*. Chesham, UK: Acumen.

Kaplan, R. S., & Norton, D. P. (1992). The Balanced Scorecard - Measures That Drive Performance. *Harvard Business Review, 70*, 71-79.

Kim, Y. G., & March, S. T. (1995). Comparing Data Modeling Formalisms *Communications of the ACM* (Vol. 38, pp. 103-115).

Landau, M. (1969). Redundancy, rationality, and the problem of duplication and overlap. *Public Administration Review, 29*(4), 346-358.
Lankhorst, M. (Ed.). (2009). *Enterprise architecture at work: Modelling, communication and analysis*. New York: Springer-Verlag.
Laudon, K. C., & Laudon, J. P. (2012). *Management information systems: managing the digital firm*. Boston: Pearson.
Leonardi, P. M. (2007). Activating the informational capabilities of information technology for organizational change. *Organization Science, 18*(5), 813-831.
Liu, B. (2010). Sentiment analysis and subjectivity. *Handbook of Natural Language Processing*, 627-666.
Liu, B. (2011). *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*: Springer.
Malachowski, A. R. (2010). *The new pragmatism*. Durham, UK: Acumen.
Mason, R. O., & Mitroff, I. (1973). Program for Research on Management Information Systems. *Management Science, 19*(5), 475-487.
Norton, D. F. (Ed.). (1999). *Hume* (Vol. 2). Cambridge, UK: Cambridge University Press.
Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems, 24*(3), 45-77.
Pennock, D. M., Lawrence, S., Giles, C. L., & Nielsen, F. A. (2001). The real power of artificial markets. *Science, 291*(5506), 987-988.
Recker, J. (2010). Continued use of process modeling grammars: the impact of individual difference factors. *European Journal of Information Systems, 19*(1), 76-92.
Recker, J. C., Rosemann, M., Indulska, M., & Green, P. (2009). Business process modeling: a comparative analysis. *Journal of the Association for Information Systems, 10*(4), 333-363.
Schawbel, D. (2009). *Me 2.0: Build a powerful brand to achieve career success*: Kaplan Publishing.
Scheer, A. W. (2000). *ARIS--business process modeling*: Springer Verlag.
Schoemaker, P. (1995). Scenario planning: a tool for strategic thinking. *Sloan Management Review*.
Schwenk, C. R. (1990). Effects of devil's advocacy and dialectical inquiry on decision making: A meta-analysis. *Organizational behavior and human decision processes, 47*(1), 161-176.
Sowa, J. F., & Zachman, J. A. (1992). Extending and formalizing the framework for information systems architecture. *IBM Syst. J., 31*(3), 590-616.
Surowiecki, J. (2005). *The wisdom of crowds*. New York: Random House Digital, Inc.
Tennent, J., & Friend, G. (2005). *Guide to business modelling* (Vol. 6): Bloomberg Press.
Turban, E., Aronson, J., & Liang, T. P. (2005). *Decision Support Systems and Intelligent Systems 7"'Edition*: Pearson Prentice Hall.
Uzgalis, W. (2010). John Locke. *The Stanford Encyclopedia of Philosophy. Stanford: Stanford University. Disponível em: http://plato. stanford. edu/archives/fall2008/entries/locke/. Acesso em, 17*.
Van Der Aalst, W., & Van Hee, K. (2004). *Workflow management: models, methods, and systems*: MIT press.
Van Der Zee, J., & De Jong, B. (1999). Alignment is not enough: integrating business and information technology management with the balanced business scorecard. *Journal of Management Information Systems*, 137-156.
Von Hippel, E. (2007). Horizontal innovation networks—by and for users. *Industrial and corporate change, 16*(2), 293-315.
Ward, J., & Peppard, J. (2002). *Strategic planning for information systems*. Chichester: Wiley.
Wijnhoven, F. (2009a). *Information management : an informing approach*. London ; New York: Routledge.
Wijnhoven, F. (2009b). *Information management: an informing approach*. London, UK: Routledge.

Wijnhoven, F. (2012). The Hegelian Inquiring System and a Critical Triangulation Tool for the Internet Information Slave: A Design Science Study. *Journal of the American Society for Information Science and Technology, 63*(6), 1168–1182.