Teaching Business Intelligence With Real Life Puzzles

Wanda Prestus
The Norwegian School of IT, wanda.presthus@nith.no

Bendik Bygstad
The Norwegian School of IT, bendik.bygstad@nith.no

Follow this and additional works at: http://aisel.aisnet.org/ukais2011

Recommended Citation
http://aisel.aisnet.org/ukais2011/40
TEACHING BUSINESS INTELLIGENCE WITH REAL LIFE PUZZLES

Wanda Presthus  
The Norwegian School of IT  
Email: wanda.presthus@nith.no

Bendik Bygstad  
The Norwegian School of IT  
Email: bendik.bygstad@nith.no

Abstract

Business Intelligence (BI) is a fashionable and popular field both in industry and academia. Still, it represents a pedagogical challenge for the lecturer. In this study we report from an introductory Bachelor course in BI and reflect on the learning process. Our focus is how to make BI education more fun and motivating for the students. We conducted a small Action Research study in a university college with a class of third year Bachelor students. Our main pedagogical approach was Puzzle-Based Learning, using real world puzzles.

Our main contribution is that the real life puzzle approach is a powerful method to teach BI concepts and processes. Real life problems are conceived by the students as interesting and motivating. Real life data create a rich and abundant source for problem solving using BI techniques. Finally, real life solutions represent an exciting target for the students to reach, with a strong feeling of satisfaction when the solution is found.

Key words: Business Intelligence, data mining, ETL, Puzzle-Based Learning

1.0 Introduction

Often, learning is associated with “no pain, no gain”. For graduate students of BI, the pain may be long lasting. BI consists of a very rational and logical approach, long sequences of actions, and endless categories of information. Add to this a plethora of tools and processes; most of them abbreviated by three letters, and you have a traditional boring subject in college. This is illustrated by a (good) BI text book in which the authors use a whole of nine abbreviations in one sentence: DSS, CRM, SCM, ERP, KM, PLM, BAM, BPM, EIS (Turban, Sharda et al. 2011), p. 105.
For many students, this may be experienced as a long wandering in the desert, before they can translate these theories and concepts into practice and pay checks. For the academic lecturer in BI, this presents a pedagogical challenge. It is not enough to assure the students that this both interesting and important in a business context. For the average BI student, business analysis may be years away while the boring class is highly present. Thus, teacher enthusiasm is not enough in order to motivate the students to learn. In this situation a shorter term strategy is needed.

Our point of departure is that having fun is an important experience in human life and undervalued feature of education (Moursund 2007). It is a sad fact of modern education that the students are often bored. Why are students bored? Often, they do not perceive what they learn to be useful: “this does not concern me”, and they do not connect what they learn with feelings or engagement.

We do not argue that the teacher needs to be a stand-up comedian; however, thinking back most people will remember something from childhood when they could have fun and learn something at the same time. For example, with only a deck of cards, a child can learn numbers, colours, shapes (hearts, spades) and even Medieval titles (kings, knaves).

Often, working with puzzles is associated with having fun. Puzzles are amusing because they are based on a mystery or real world problem; you know that there is one correct answer, and that there is rational reasoning behind the solution. Potentially, all these aspects are present in BI problem solving. Therefore, in this paper, we investigate this potential, and our research question is: How can a Puzzle-Based approach contribute to improved learning in BI?

Our empirical evidence is a course of BI at Bachelor level. After reviewing the relevant literature, we describe in some detail four real life puzzles given to the BI students. In our discussion section, we theorise on our findings, and discuss the viability and limitations of Puzzle-Based teaching.

2.0 Literature Review
In this section, we first present the basics of BI, and highlight some pedagogical challenges. Then we review the concepts of Problem-Based Learning and Puzzle-Based Learning, and compare them to other learning strategies. Finally, we present our framework for analysis.
2.1 Business Intelligence As An Industry Discipline

Graduate students are being taught that Business Intelligence is nothing new, but used to be called decision support or executive information systems. Today, it comprises various tools, applications, processes, databases and architectures for all levels of an organisation by providing access to (often real time) data (Turban, Sharda et al. 2011). The major aim of BI is to turn a company’s data into actionable information. The process BI is based on Simon’s framework for decision making (Simon 1977) and consists of four steps: gathering data, turning data to information, making decisions and finally take action. Research has shown that many companies do not make it through the whole BI process, and tend to halt after having achieved information, but not taking action (Davenport, Harris et al. 2001; Overby, Bharadwaj et al. 2006; Howson 2008). Therefore, teaching BI should also train the students to pay attention to not only data analysis, but also on how to make decisions based on the data.

The BI kit comprises spread sheets, data warehousing, and dashboards, as well as applications for queries, online analytical processing, and data mining. However, non-technical tools such as your head, pen and paper and maps may also be used. Modern BI tools can also process all kinds of data: structured, semi-structured, and unstructured, as well as data appearing as numbers, text, web pages, sound, pictures, or film, as shown in figure 1 below (Business Objects 2007).

![Figure 1: Various types of data and sources (Business Objects, 2007)]
Authors of well-known BI- and database books, (Moss and Atre 2003; Connolly and Begg 2010), all focus on the importance of the ETL process. ETL stands for Extraction, Transformation and Load, and is a key concept in any data warehouse or database. A data warehouse is “…a collection of data in support of management’s decision making...” (Connolly and Begg 2010, p. 1147). The authors claim that the major reason for BI projects to fail is neglecting the ETL process, accompanied by Turban et al who write that the ETL process “typically consumes 70 percent of the time in a data-centric project” (Turban, Sharda et al. 2011, p. 433).

Data mining means searching for (usually undefined) patterns in large amounts of data. Mining is divided into somewhat various categories, but usually comprises the ones showed in table 1, as described in Moss and Atre (2003).

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Association</td>
<td>If the customer purchases airline tickets for the whole family, there is X% chance for car rental also</td>
</tr>
<tr>
<td>Sequence</td>
<td>If the customer purchases a washing machine, there is X% chance for purchase of a dryer within 6 months</td>
</tr>
<tr>
<td>Classification</td>
<td>Which customers who are likely to purchase a product (predefined group)</td>
</tr>
<tr>
<td>Clustering</td>
<td>The group is divided after the mining, i.e. according to faulty products</td>
</tr>
<tr>
<td>Prediction</td>
<td>Based on historical data, such as stock marked, weather forecast</td>
</tr>
</tbody>
</table>

**Table 1: Various forms of data mining**

Perhaps the most famous example of data mining is the grocery store Wall Mart (Turban, Sharda et al. 2011). By means of association (explained in table 1), Wall Mart could find out not only “On which week day do we sell most cans of beer”, but also: “Which products do our customers purchase together with beer on which week days and by what gender”. The answer read: “beer and diapers on Thursdays and Saturdays by male customers”. Wall Mart successfully used this information and decided that they did not need to have both products on sale on these days, and they also placed the two products together in the store.

In addition to the inceptive data mining, the new trends in today’s BI are text- and web mining (Turban, Sharda et al. 2011). Text mining has the same purpose as its forefather data mining, but manages unstructured and semi-structured data from text documents such as Microsoft Word, e-mail, social media and more, as illustrated in figure 1 above. Web mining was initially applied by Etzioni, and refers to the extraction of relevant
information from Web pages (Turban, Sharda et al. 2011). Web- and text mining can serve as foundations for sentiment analysis, a technique used to identify positive or negative opinions towards a product or service, usually expressed by customers. Sentiment analysis can be performed by a number of applications, such as SAP Text Analysis or IBM’s SPSS, or it can be performed manually, which would of course be time consuming, but less expensive.

2.2 Business Intelligence As An Academic Topic

Decision support has been taught as an academic topic for at least 30 years (Power 2007), often with a mathematical foundation. Business Intelligence is now being taught at a growing number of universities and colleges around the world. While there are relatively few studies on teaching BI, reports and articles are now beginning to emerge. In a blog from Manila with the timely title “What I learned teaching Business Intelligence to Undergrads”, Yapjoco reported that he used the tool QlikView to teach Business Intelligence for one semester and how he struggled to make the students develop analytical skills as well as an understanding of business value. His conclusion was that teaching BI can only be a success if the students have experienced relevant tools and solutions (Yapjoco 2010). In Gothenburg, at the Centre for Business Solutions, Ask et al reported on a programme called SANTE (Scandinavian Academic Network for Teaching Enterprise Systems). Students from various Swedish universities competed on solving a task while given real life accounts from a real manufacturing company (Ask, Magnusson et al. 2009).

On a more technical level, a particular interesting study was published by Fang & Tuladhar who described the concepts of data warehousing, OLAP and their experiences of successfully teaching this to students (Fang and Tuladhar 2006). Their paper was cited by Mrdalj who concluded that it is important to combine theory and practise, as well as to integrate real-world examples in the lectures. This “...made this class exciting what otherwise used to be dry presentations of the analytical data mining techniques” (Mrdalj 2007, p. 38.) Mrdalj advised using a large variety of data samples to explore. He also pointed out that preparing for such a course was extremely time-consuming for the lecturer.
2.3 Problem-Based Learning Versus Puzzle-Based Learning

In the widely cited framework of Dunn and Dunn, a total of twenty learning styles is described, including analytical skills, time of day, working in groups, and lighting system (Dunn and Dunn 2003). Not one of these elements holds the name of having fun!

An alternative approach is Problem-Based Learning (Savery 2006). This approach is about creative thinking on real life problems, and has open-ended solutions. The teacher facilitates the students’ learning and reflection while the students work in groups. Problem-Based Learning has more focus on the process – on how the students learn – and less on the outcome; usually there are multiple correct answers. Thus, exercises where Problem-Based Learning may be useful could read something like: “Describe how a company can benefit from web mining”.

While Problem-Based Learning does not provide one correct answer, a puzzle will always have one fixed solution. Puzzle-Based Learning was described by Michaelewicz and Michaelewicz, who showed how it can be used to teach engineering students to solve problems. The point is to make engineering students think more creatively. They believe that there are multiple reasons behind most students’ enthusiasm for the Puzzle-Based Learning. Puzzles are educational, thought-provoking, and make problem-solving entertaining. In addition, puzzles can illustrate different disciplines, such as statistics, and application areas, such as finance, in a simple way. The conclusion of the puzzle can be transferred to solving real-world problems (Michaelewicz and Michaelewicz 2007).

Well known examples of puzzles are crosswords, jigsaw, and solitaire (Moursund 2007). Some puzzles can only be solved by thinking outside the box, such as the problem with the farmer having to cross the river with his vegetables, chicken and wolf with only room for one item at the time in his boat. How does the farmer prevent the chicken eating the vegetables, and the wolf eating the chicken? The answer to this puzzle requires some out-of-the-box thinking: bring the chicken over first, because the wolf will not eat the vegetables. Go back in the boat alone, to fetch the wolf. Bring the chicken back again, and fetch the vegetables. Finally, go back alone and get the chicken. This brainteaser dates back to AD 732 (Michaelewicz and Michaelewicz 2007).

In the same manner, Higley and Marianno focused on the fun factor in teaching their engineering students; they constructed a mystery of a fictitious murder (Highley and
Marianno 2001). Their engineering students were required to solve this murder using “nuclear techniques”, such as gathering evidence. The students built knowledge through interviews and reasoning. The authors also commented:

“This mode of teaching requires extensive hands-on faculty participation. However, the potential long-term benefit is increased comprehension of course content as well as greater student interest and retention.” (Higley and Marianno, 2001, p. 105)

We suggest that the principles of Problem-Based Learning (real life problems) and the Puzzle-Based Learning principles (thinking outside the box) are relevant for making BI education more fun and motivating for the students. Our approach was to take a real life situation, and make a puzzle from it. In the next section, we describe this framework.

2.3 A Framework For Analysis: Real Life Puzzle

Summing up this review, there are three aspects associated with Problem-Based Learning and Puzzle-Based Learning in a BI context that we find significant.

Real life problems: According to the Problem-Based Learning, a problem should be ill-structured and come from the real world, because the real world is often messy, which motivates the students better to find a solution (Savery 2006). We agree that the problem should come from the real world, but it does not need to be unstructured. We believe that BI students are more motivated by real life problems than by theoretical problems. While students of mathematics and science are fascinated by constructed puzzles (Highley and Marianno 2001), our view is that BI students are primarily motivated by real life puzzles, as these students are more business value oriented, like the case with Wall Mart described in section 2.1.

Real life data: Second, BI is characterised by large volumes of real life data, for example, structured and transactional data from order and sales, and unstructured data from social media and Word documents (see figure 1). Extracting patterns out of large data amounts is a key concept within BI and is conducted through ETL and data mining (Turban, Sharda et al. 2011). This represents a unique opportunity of puzzle investigation
in huge volumes of data. In the SANTE programme, real accounts are provided from a real world company (Ask, Magnusson et al. 2009).

**Real life solutions:** Lastly, Puzzle-Based Learning emphasises that the fun of puzzles is determined by the fact that there is only one correct solution. The correct solution is often surprising, but completely logical according the data material. In Problem-Based Learning, the activities must also have value for the real world (Savery 2006) however; there is not one rigid solution. Business Intelligence has a logical approach and its main issue is about making decisions based on data (Davenport, Harris et al. 2001). Data in real life may not always provide a clear pattern, even by means of the plethora of BI tools existing today. However, we believe that in a learning situation, there should be one revealing answer so that the students will feel satisfaction having spent quite some time solving the puzzle.

These three criteria (real life problems, real life data, and real life solutions) will be used in our analysis and discussion.

### 3.0 Method

The research approach was a teaching case, conducted as a small Action Research project. Action research is a recognized approach in several fields, such as organisation development, innovation, health care and education (Reason and Bradbury 2007) and in Information Systems research (Baskerville 1999). Recently, there has been an increasing interest in using action research in higher education (Norton 2009). There are multiple versions of Action Research, and we chose the Canonical Action Research, because it is iterative, rigorous and collaborative (Davison, Martinsons et al. 2004).

The teaching case was carried out at the Norwegian School of IT during the autumn semester of 2010. The class counted 28 students, attending the ten ECTS Business Intelligence course, with was mandatory in the fifth semester of the Bachelor of Information Technology. The learning objectives of the BI course were: master the key concepts; acquire knowledge of the business value and ethical issues; apply theories on practical problems; and handle an assortment of tools such as a data warehouse and reporting and analysis, as well as spreadsheet. Following the principles of Canonical Action Research cycle (Davison, Martinsons et al. 2004), the following steps were conducted:
Diagnosis: This was an introductory course in BI with certain pedagogical challenges. First, a number of terms and abbreviations was needed to make a conceptual fundament. Second, the students may struggle to understand the business value of BI. Searching for new ways to make BI learning more fun, we reviewed the literature for possible approaches. Our framework, based on Puzzle-Based Learning, was derived from the literature as described above.

Action Planning: The first author of this paper designed a pedagogical approach drawing on Puzzle-Based Learning. Four puzzles, each of them illustrating a key BI concept and business value, were constructed and documented for the use in the class.

Action Taking: One puzzle was conducted for each lecture, four weeks in a row. Further details are described in section 4.

Evaluation: Observation during the lecture was used to understand how the students solved the puzzles. Then, an anonymous survey on paper was conducted with all students to measure their perceived learning and satisfaction. They were asked whether the puzzles were fun, and if the puzzles helped them understand key concepts of Business Intelligence. They were also informed that the aim was to produce research similar to Fang and Tuladhar's paper, which they had previously read. Such an agreement “contributes to the internal validity of the research” (Davison, Martinsons et al. 2004, p. 69). Moreover, the authors suggest multiple sources of data to strengthen the validity.

Specifying Learning: This stage is about reflecting on the outcomes so far. The outcomes of the actions were carefully analysed and assessed practically and theoretically, drawing on related literature. We analysed in detail the four puzzles, and how they were perceived by the students, and also the perceived differences between the four puzzles. Then we drew on some theoretical concepts from game and puzzle research. The main results are presented in this paper.

This paper reports on one loop in the Action Research cycle. Deciding whether to exit or go one more iteration should be related to whether the goal has been reached, and needs to be described in the report (Davison, Martinsons et al. 2004). As this was a teaching situation, it was decided in advance to do only one loop conducting four puzzles. Previous experience indicated that this was sufficient, also taking into consideration the time-consuming task for the teacher of finding such real life puzzles. Also, going one more
loop with the same students and the same puzzles would be pointless, as they would know the answer to the puzzles.

4.0 Four Puzzles In BI Teaching And Their Outcomes

In this section, we present how the puzzles were given to the students, how the students attempted to solve them, and finally the results of the students’ effort.

4.1 Puzzle One: Rusting Meters In London.

The objective of this puzzle was to introduce the concept of data warehouse. The whole class solved the puzzle as one group. They had only two tools; their head and making queries to several simulated databases, residing in the head of the teacher.

**Puzzle:** The idea for this task was taken from a paper by Foss and Bond, describing how parking meters were mysteriously rusting in a confined area of London (Foss and Bond 2005). The problem was that the meters were rusting from within, and only in a given borough in London. Replacing them was expensive for the community. How could this problem be solved?

**Process:** The students were told that they had access to all possible information on the city. They were explained that this meant querying the lecturer, who would act as several databases.

The students immediately started guessing, forgetting any strategy for solving the problem. Typical random guesses were: “The rusty meters are shorter, or made of different material!” They had to be reminded of making use of data, which could be extracted from the lecturer. Slowly, the nature of questions changed. For example, they started to inquire about the income of the people living in the streets with rusty meters, and received the answer: “They make less money.”

**Students:** “Do they wash their cars with hoses, because they cannot afford automatic car wash?”

**Lecturer:** “No, they do not have hoses. But they cannot afford automatic car wash, either.”

**Students:** “Do we know where they shop?”

**Lecturer:** “Yes, in low-price stores.”

**Students:** “What do they buy?”
Lecturer: “Many plastic items, for storing food. In all shapes and forms.”

The lecturer (like any database) would not respond if the query was incorrect:
Students: “It has to be some kind of vandalism!”
Lecturer: “Your query does not return any data. Please rephrase your query.”
Students: “OK. Do the rusty meters show sign of tampering when examined?”

Solution: It took about one hour of queries and answers before one student suggested something with ice cubes. This was indeed the correct solutions: the car owners in the streets with rusty meters made “ice coins” which would later melt inside the meter! The class also discussed how the authorities acted on this information.

4.2 Puzzle Two: Dr. Snows “Ghost Map” From 1854.

The next puzzle was “Business Intelligence anno 1854”, and showed that data can come in many shapes. It was also given to illustrate how people tend to make decisions not on information, but on previous experience or dogmas. In 1853-4 there was an epidemic in London taking 10,738 lives. The medical profession, authorities and waterworks all agreed that cholera infects though “bad air”. Dr. Snow eventually suspected the bacteria to spread through water, and went from door to door in Soho, collecting data on his “Ghost Map”, shown in figure 2 (Johnson 2010).

Figure 2: Dr. Snow’s Ghost map (Johnson, 2010)
Puzzle: The ghost map with red and blue dots was shown to the students: what does it show, and how to act on the data?

Process: Like in puzzle 1, the students acted as one group, and could only use their head and the lecturer as omniscient. They were not allowed to search the Internet.

Lecturer: “Has anyone seen this map before?” (No student had seen it before.)
“Where is it?” (Soon someone guessed London.) “How old is the map?” (Someone suggested fifty years old, then after 3-4 attempts: 1850’s.)

Knowing that the map was over one hundred years old, they quickly matched the blue dots with water. Then they started their quest for the red dots. “Horses?” – but quickly landed on people. The students had clearly learned from lecture one. Questions where now of the character: “Are the people connected to the workhouse? Do they work at the brewery? Is it the map of a land owner, visualising tenants who are late with rent?”

Solution: One student eventually connected the map with cholera outbreaks. Having revealed that the map indeed indicated death by cholera, they immediately matched the death having to do with the one water pump with the numerous red dots drawn to it in “Broad Street”.

Then they were challenged: Having turned data into information “death is caused by water, not by air”, how to act on this information? The decision seemed obvious to the students: “Close down that pump”. However, few students guessed the incredible half decade it actually took for the authorities to act correctly on the new information.
The students quickly wanted to know whether Dr. Snow lived to receive credit for his discovery – which he sadly did not. It would take almost 50 years for this message to be believed and acted upon (Wikipedia, 2010). As the drawing reads in figure 3: “Open to the poor, gratis, with permission of the parish” (Nevolution 2009).

4.3 Puzzle Three: Text Mining.

As presented in the literature review, web- and text mining is increasingly popular (Turban, Sharda et al. 2011), which can make a foundation for sentiment analysis. During the ash crisis in spring 2010, the airline companies Norwegian and SAS used Facebook as “Agile CRM” (Presthus and Bygstad 2010). Prior to the exercise, by means of a self-developed web content mining tool, all words from Norwegian and SAS pages on Facebook had been extracted.

**Puzzle:** The students were asked to categorise the data into positive and negative words, and accumulate each category, in order to find out whether the atmosphere was negative or positive.

**Process:** This time the class was divided into five groups, ranging from two to five students in each group. Their tools was again their head for manually categorising the words, Excel for making graphs, and over 60 000 words extracted from two airline companies from Facebook in an Excel file. At first, the students were overwhelmed by
the amount of words. The teacher needed to motivate by telling that this task had been performed already, and had taken less than two hours. Attention was also drawn to the fact that this was real data from real web sites.

**Solution:** The students in groups spent over one hour each, and even insisted on more time to finish the task. The result was five graphs from the students, all showing that the highest graph would be negative words by Norwegian. One example is presented in figure 4 below.

![Figure 4: An example of graph by the students having completed puzzle 3](image)

All students also reported the text analysis to be categorisation, as we had a predefined set of categories: positive and negative words. They also correctly identified the mining as web content mining (as opposed to web usage or web structure mining).

### 4.4 Puzzle four: ETL/Data mining.

The learning objective of the last puzzle was to understand the ETL (Extraction, Transformation, and Load) process, and to conduct numerical data mining.

Like puzzle 3, the students were placed in groups, and given two files with numeric data from Telenor’s Sales and Marketing department. The task would read: “one file shows how many customers who received SMS on their mobile offering to purchase Rybak’s singel “Fairytale” (extracted from the Marketing department). The other file shows how many customers who actually purchased “Fairytale” regardless of having received an offer or not (extracted from Sales department)”.

**Puzzle:** The management wants to know how effective the campaign was, that is: how many who received an offer on their mobile, also purchased the single “Fairytale”? 
**Process:** The two files consisted of (modified) customer numbers. The one file, Offer, contained 100,000 numbers, the other, Sale, contained 5239. Moreover, the latter contained data in different format, as illustrated in figure 5 below. Before the students could compare the data, they had to somehow remove the prefixes “1-”. (The correct answer: 2666.) Two customers had also purchased the song 4 times each.

![Figure 5: Two files, “extracted” from Marketing and Sales Department](image)

**Solution:** Three groups presented the correct answer, while most groups presented a nearly correct answer. The latter managed to transform the data and compare the two source files, thus extracting the matching instances, but failed to identify the two customers having purchased four songs each. When confronted it was not quite correct, one student shrugged and replied: “How important can it be?”

The table below summarizes the four puzzles.

<table>
<thead>
<tr>
<th>Puzzle</th>
<th>Key BI concepts</th>
<th>Learning objective</th>
<th>Applied BI technique</th>
<th>Outcome by students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rusty meters</td>
<td>Data warehouse</td>
<td>Collecting data from several sources in order to solve a problem</td>
<td>Simulating queries from multiple databases</td>
<td>Found the answer within one hour</td>
</tr>
<tr>
<td>Dr Snow’s Ghost Map</td>
<td>Data mining, the BI process</td>
<td>Map can be data source. Importance of acting upon information</td>
<td>Discover patterns. Decision making</td>
<td>Found the answer within one hour</td>
</tr>
<tr>
<td>Norwegian and SAS on Facebook</td>
<td>Text mining, web content mining</td>
<td>Sentiment analysis is “hot”. Experience tools</td>
<td>Categorisation. Visualisation</td>
<td>All groups reached similar results</td>
</tr>
<tr>
<td>Effect of campaign at Telenor</td>
<td>ETL process</td>
<td>Data sources often have different formats</td>
<td>Transform data. Match unique numbers appearing in both files</td>
<td>Only three groups reached the correct answer</td>
</tr>
</tbody>
</table>

**Table 2: Summing up the 4 puzzles**
The four puzzles were all based on a real life situation, using real life data. All students found the answer matching the real life solution in all puzzles, except in puzzle number 4, where the majority was close to finding the correct answer. The next section discusses these findings.

5.0 Discussion
We begin by returning to our research question; how can a Puzzle-Based approach contribute to improved learning in BI? Our discussion follows the framework presented in the research review. We start by discussing the use of real life problems, proceed with discussing the use of real life data and then discuss real life solutions. Finally, we assess the validity of our findings.

Our point of departure was than the fun factor is underrated in much teaching, and the puzzles were designed in order to combine fun and learning. Table 3 summarizes the students’ replies to the question: “Was the puzzle fun?”

<table>
<thead>
<tr>
<th>Fun?</th>
<th>Yes</th>
<th>Somewhat</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Puzzle 1</td>
<td>15</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Puzzle 2</td>
<td>16</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Puzzle 3</td>
<td>11</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Puzzle 4</td>
<td>8</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Sum</td>
<td>50</td>
<td>16</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 3: The fun factor of each puzzle

As table 3 shows, the students found all of the puzzles to be quite fun. The most popular puzzles were number 1 and 2, while puzzles 3 and 4 were perceived somewhat less fun. From the comments of the students, we note that some of the students reacted to the required amount of work. We proceed with elaborating what made the puzzles exciting, but also pedagogically successful.

5.1 Real Life Problems: Intriguing And Dramatic
We found that the students’ enthusiasm was associated with interesting real life problems. As described in section 4 above, the students asked many questions on the facts of the cases, and generally engaged in the puzzles. We note, however, that the students were more engaged in the three first tasks. According to the survey, and the results in class, the three first tasks were solved successfully, whereas the fourth was only completed by a few groups. The students were very interested in the outcome of Dr. Snow’s discovery,
but less about Telenor’s campaign. The deaths of ten thousand people in 19th century London was more intriguing than a relatively trivial business problem at a telecom company. As students reflect in the survey:

“Interesting to see how long the government took to accept Snow was correct”

“Fun task; as it is from real life”

Most research on how to teach engineering leans on constructed problems. Examples are murder mysteries (Highley and Marianno 2001), or river crossing (Michaelewicz and Michaelewicz 2007; Moursund 2007). Moursund argues that it is up to the students to make the connection between the constructed examples and the real world. Although young students will struggle to make such connections, a good teacher can provide the link, and then the student can draw on this experience later.

Our evidence points to another direction, because particularly business students tend to perceive real world problems as more motivating. We note that Mrdalj’s data mining students reported that it was the real life problems which made the course interesting. Just like our subject, Mrdalj’s course was more business oriented than technical; the students were given a business problem to solve, such as customer marketing (Mrdalj 2007).

5.2 Real Life Data: Genuine And Large Volume

Our findings reveal that data from real life was motivating. The volume of data was modest in the two first cases, while quite substantial in case 3 and 4. Working with real life data challenged the students’ creative and analytical abilities, and made them very curious to find the answer. Two typical student comments were:

“Challenged creativity and logical thinking”

“Jumpstarted thoughts and became curious of the reason”

In line with this finding, we noted that case 4, modified telephone numbers from Telenor, was perceived somewhat less fun. We assume that the reason was that the data consisted only of thousands of numeric data, and it hard to see the business value. As several students write in the survey:

“Useful, but time consuming. Would be better to finish together after 10 minutes”

“Tiresome, too much data”
How much data is needed to conduct an interesting analysis? Mrdalj’s students liked real life data, but wanted a larger amount (Mrdalj 2007). Our experience is that the volume of data in itself is not crucial for how interesting the students find the exercise. Rather, it is the patterns of the analysed data which contribute to solve the real world problem. Both puzzle 2 and 3 were perceived as equally fun, although the amounts of data were very different. In the case of Dr Snow, there were only one map with dots, and the text mining case had over 60 000 words to process. In these two cases, data had been turned into information, whereas in puzzle 4, the numbers were only raw, primitive data presented in two txt-files, as shown in figure 5 above.

A key feature is that Business Intelligence is based on certain logic in querying; you cannot ask directly for the solution, but you have to frame your questions following certain techniques related to the available data. It is the pattern of the data that will reveal the solution. These patterns may be surprising: a map with streets and dots actually lead to a new discovery on how cholera spreads.

Real life data can be found on web pages (for example Eurostat), by asking companies for data, or taking a documented case from the real world and creating a puzzle of it. There are also challenges with real world data. One challenge is that it puts a lot of pressure on the lecturer in order to supply these data, and the other is ethical issues associated with privacy, business property and secrecy. We will return to these issues in section 5.5.

5.3 Real Life Solutions: One Solution, Known To Exist

Having been presented the puzzle, the students were quick to ask whether the answer existed, and whether there was only one solution. We find that the students should be told in advance that there is one correct solution to the problem, and that the answer can be derived only from the data. Moreover, the solution needs to be intuitive, and have an impact on business value. Three students commented:

“It made us understand how to reach a solution by analysing a database, and reach a conclusion”

“Fun to reach answer, it was not obvious”

“Very good to learn how this information can be used to make money”
In contrast, the Telenor case was less motivating; even though the students knew there was one correct answer (“2666”), some students failed parts of the ETL process and landed on 2672. When confronted that it was incorrect, some students shrugged and said “Isn’t it close enough”?

Perhaps the puzzle would be more motivating had the consequences of the solution been more radical, such as giving a large bonus or gift to the detected customers. This was not the case at Telenor, and may be one drawback of using real life puzzles: real life consists of many mundane cases and situations.

Previous research has shown that there must be rules in order to solve the puzzle (Highley and Marianno 2001). In our case, the students were not allowed to search the Internet, and they were informed about these rules in advance. None of the students broke any rules. Although Michalewicz and Michalewicz warn that it may take some time to reach the solution, we still believe that the solution should not be fabricated when using real life data. Moreover, the solution must be perfectly logical, based on the evidence. Having spent quite some time finding the correct answer, the students are not motivated for another fifteen minutes for the answer to be explained.

5.4 Too Much Fun?

However, can it be too much fun? Can there be too many puzzles in the teaching? Is there a risk that the students become passive in the sense that they stop reading themselves, and expect the learning to be wrapped in entertainment? Last year, when evaluating the course using an online survey, one student complained about too many “parlour games” (named by the student him-/herself) and according to this year’s survey on paper, some students found the puzzles taking too long:

“Spent too much time, we would have got the idea with less words”.

“Overall, too much time spent. Less interesting whether fun, should focus on useful.”

“Would like to work more with SAPBW, and spend less time on concepts and abbreviations. Other than this, this subject is good 😊”

This may indicate that the students see the brainteasers for what they are meant to be, namely a tool for learning BI concepts and processes. Interestingly, Higley and Marianno
report that the *teachers* found the puzzle to be too much fun, and were giving the students too many incoherent clues. This brings us back to the introduction where we contend that the teachers’ enthusiasm is not enough.

### 5.4 Summing Up

Overall, we argue that the real life approach is a powerful approach to learn BI. *Real life problems* are conceived by the students as interesting and motivating. *Real life data* also create a rich and abundant source for problem solving using BI techniques. *Real life solutions* represent an exciting target for the students to reach with a strong feeling of satisfaction when the solution is found. Moreover, the real life approach teaches the students to think “outside the box”.

This is illustrated with a number of characteristic comments from the students:

> “Very fun way to make us think differently”

> “I would probably be clueless without this puzzle”

There are however limitations and challenges to this type of learning approach. We will briefly assess them in the next section.

### 5.5 Limitations

First, we have to acknowledge that this was a small study, consisting of a one semester course with some twenty students present. We still believe that our findings are consistent with previous research, and valid for other contexts.

Second, previous research (Mrdalj 2007) has emphasised that it is time consuming and difficult to identify such real life puzzles. Puzzles such as Rusting Meters and Ghost Map may only be used once for each cohort – even provided that no student has heard of it before somehow. The BI lecturer runs the risk of running out of puzzles from the real world. At the moment, there is no “World Puzzle Bank” although it should be possible to draw on other peoples puzzles.

On a more technical level, when Mrdalj warns that making a course with real life data is time-consuming, today’s BI techniques such as ETL can extract large amounts of real life data from multiple sources such as Facebook and other web pages. On the other hand,
Teaching Business Intelligence With Real Life Puzzles

ethical issues may prevent use of real life data, as was the case of Telenor. It would not be good practise to apply peoples personal data extracted from real registers.

6.0 Conclusion
In this paper we investigated how to teach some of the core concepts and practises of Business Intelligence at a university college, while at the same time making learning fun. Students were given four tasks, covering data warehouse, ETL process, data- and text mining, using various BI tools such as their head, a map, and Excel.

Our conclusion is that real life puzzles is a powerful method to teach BI concepts and processes. Real life problems are interesting for the students, real life data create a rich source for problem solving using BI techniques, and real life solutions are motivating.

There are also limitations associated with this approach. Obviously, Puzzle-Based Learning does not replace the usual classroom teaching, but serves as a supplement. Moreover, preparing puzzles represents a considerable burden for the lecturer. Collecting a large amount of data may be challenging, and ethical issues have to be addressed. Finally, some puzzles may only be used once for each class.

Further research on Puzzle-Based Learning can be conducted in other subjects than BI. Within the BI field, some related venues of research could be investigated. For example, a comparative study of constructed BI puzzles could give additional insights in the student learning process. Further research could investigate whether puzzles can affect the learning curve – will the students remember better by solving a puzzle?

References


