Requirements specification for Sale business process

1. General description of business process

a. A general description of the business process and a description of the performance metrics generated by this process, possible current analytical problems.

The process of book selling is as follows: the customer enters the bookstore and chooses a book or books that she/he would like to buy. The buyer goes to the one of cash desks and gives the book (or books) she/he has chosen. The barcode of the given book is read by the system, to which the seller is logged in. In this system the barcode is connected to the book ISBN, its title and price. The sales system generates a bill on which the price for each book is specified. The customer can buy several copies of the books, what is also denoted on the bill. Then the client pays the bill. In Book's bookstores it can be paid by cash, by card or by check.

The increase in the monthly profit of individual bookstores at a level not less than 0.5% monthly compared to the previous month.
The increase in the number of copies sold per month at a level not less than 0.5% monthly compared to the previous month.

b. Typical questions

Compare the number of copies of sold books in terms of their genres.
What are the most popular authors in this month?
Give the best-selling titles in this month.
Compare the profit of the current month with the profit of the previous month.
Compare individual sellers in terms of the number of books they sell.
How many of the most widely read titles sold individual sellers?
Give the average monthly profit for the last year.
Give a total profit for the last year.
Compare the profit from the last month with the average annual profit.
Give profit from the sale of books during the holiday periods.
Specify in how many cases the seller's opinion affected the purchase of the book.
What genres of books are purchased when the seller advises?
What are the profits for individual bookstores?

c. Data

All sales data are extracted from the sales system – "BillMaster". Sales system stores information about the seller making the transaction (thanks to logging in), the transaction number, the books purchased within this transaction and the price for which the book was purchased. In addition, it is known that the margin of the store for each copy is 7%. Moreover the payment method (card, cash or check) is settled. In addition, data about employees and bookstores are stored in the EXCEL sheet.

2. Data sources structures

BillMaster

TABLE NAME	ATTRIBUTE	ATTRIBUTE TYPE		DESCRIPTION
BOOK	A title identified by its ISBN number – specific issue and not single item.			
	ISBN	String – characters	20	PK
	Title	String – characters	30	Book title
	Genre	String – characters	20	Book genre. Allowed values: encyclopedia, album, fantasy, other, informatics, magazine, history, language, cooking, drama, poetry, thriller
AUTHOR	Book author identified by first name, middle name and surname			
	IDAuthor	Numeric		PK, auto generated starting from 1 with increment 1
	Name1	String – characters	20	First name
	Name2	String – characters	20	Middle name, by default an empty string
	Surname	String – characters	30	Surname
AUTHORSHIP	Implementation of many-to-many relationship between Author and Book. It means that a given author has written a given book. Identified by foreign keys to tables Book and Author.			
	FK_Book	String – characters	20	FK, part of PK
	FK_Author	Numeric		FK, part of PK
BOOKSTORE	A specific bookstore from the Book's Bookstore network is identified identification number.			
	IdentificationNumber	Numerical		PK – identification number
	Name	String – characters	25	Bookstore name
SALESPERSON	The salesperson in the Book's Bookstore network identified by Personal Identification Number - PIN.			
	PIN	String – characters	11	PK
	Name	String – characters	15	Salesperson's name
	Surname	String – characters	30	Salesperson's surname

BILL	A bill issued in one of the Book's Bookstores.			
	BillNumber	String – characters	15	PK – Bill identification number
	IssueDate	DateTime		Bill issue date and time (with a second precision).
	PaymentDate	DateTime		Bill payment date and time (with a second precision).
	Place	String – characters	50	Cash desk location.
	Payment	String – characters	15	Payment type.
	FK_Salesperson	String – characters	11	FK pointing at the seller issuing a bill. Implementation of the n to one relationship between Bill and Salesperson (issued by).
	FK_Bookstore	Numerical		FK pointing at the bookstore, in which the bill was issued. Implementation of the n to one relationship between Bill and Bookstore (issued in).
BOOKSALE	Implementation of n to n relationship between Book and Bill meaning a sale of the book within the bill. It is identified by two foreign keys of Book and Bill tables.			
	FK_Book	String – characters	20	FK Book, the part of PK
	FK_Bill	String – characters	15	FK Bill, the part of PK
	Price	Decimal (two digits precision)		Price in EUR (with the cent precision)
	NumberOfCopies	Numerical		Number of copies sold within the bill.

CEO Excel

Sheet 1 (Information about bookstores in Book's Bookstore network, each line describes one bookstore, line 1 is a header row):

   Column A - Bookstore identification number (numeric, 0 decimal precision),
   Column B - Bookstore name (text),
   Column C - Street and house number (text),
   Column D - Postal code (text),
   Column E - City (text).

Note: If the address changes, the data in the sheet are updated. There is always one row for one bookstore in the sheet.

Sheet 2 (Information about employees hired in the Book's Bookstore network, each line describes one employee, row 1 is a header row):

Column A - Identification number of the bookstore in which the employee is employed (numeric, 0 decimal precision),

Column B - Employees' PIN (PIN number),

Column C - Employee's name (text),

Column D - Employee's surname (text), in case of surname change the column is updated,

Column E - Date of birth (in format year - month – day e.g. 2013-12-09),

Column F - Education (text), in case of change the column is updated,

Column G - Position (text), each row means employment of a given person on a given position; in case of position change the date of end of work on the current position is written and there is a new line with a new position and a new date of employment; there are two available positions: salesperson and director, each bookstore has exactly one director,

Column H - Date of acceptance for the position (Date in format year - month - day, e.g. 2013-12-09),

Column I - Date of end of work on the current position (Date in format year - month - day, e.g. 2013-12-09), it is not set if the employee currently works on a given position.

3. Scenarios of analytical problems

Why was there an increase / decrease in sales this month?

1. Compare the number of copies of books of different genres sold in the analyzed month relative to previous months.
2. Compare the book sales at the near-the-holidays days in current and previous month.
3. What are the best-selling books in current and the previous month?
4. Compare the profits from book sales for each seller in current and previous month.
5. What are the best-selling authors in current and the previous month?
6. Compare the profits from different bookstores in current and the previous month.
7. What are the book sales in relation to the employee's seniority in a given position?
8. Compare the book sales in relation to the bookstore size, defined as a number employees employed, in current and the previous month.
9. What is the profit in relation to bookstore's distance to shopping mall?

What is the effect of advices given by shop assistants on the sales size?

1. In how many cases the seller's advice resulted in the purchase of the book?
2. What kinds of books are bought when the seller advises?
3. What titles are promoted by the sellers and what is the profit from these titles?

4. Data needed for analytical problems

Analytical problem: "Why was there a increase / decrease in sales this month?"

1. Compare the number of copies of books of different genres sold in the analyzed month relative to previous months.
   - **number of copies of books sold** - *BillMaster*, table *Booksale*, column *NumberOfCopies*

- **book genre** - *BookMaster*, table *Book*, column *Genre*
- **month of sale** - *BookMaster*, table *Bill*, column of the *IssueDate*

2. Compare the book sales at the near-the-holidays days in current and previous month.
   - **number of copies of books sold** - *BookMaster*, table *Booksale*, column *NumberOfCopies*
   - ***near-the-holidays days*** must be collected from some publicly available calendar, e.g. Google calendar

3. What are the best-selling books in current and the previous month?
   - **number of copies of books sold** - *BookMaster*, table *Booksale*, column *NumberOfCopies*
   - ***sale month*** – *BookMaster*, table *Bill*, column *IssueDate*

4. Compare the profits from book sales for each seller in current and previous month.
   - ***profit*** – calculated from selling price from *BookMaster*, table *Booksale*, column *Price*. According to the interview, Mr. Book's margin is 7%, so *profit = Price*/1.07
   - ***seller name*** – *BookMaster*, table *Salesperson*, columns *Name* and *Surname*
   - ***sale month*** – *BookMaster*, table *Bill*, column *IssueDate*

5. What are the best-selling authors in current and the previous month?
   - ***author name*** – *BookMaster*, tables *Book*, *Author*, *Authorship*; name and surname to be taken from table *Author*, columns *Imie1* and *Surname*
   - **number of copies of books sold** - *BookMaster*, table *Booksale*, column *NumberOfCopies*

6. Compare the profits from different bookstores in current and the previous month.
   - ***profit*** – calculated from selling price from *BookMaster*, table *Booksale*, column *Price*. According to the interview, Mr. Book's margin is 7%, so *profit = Price*/1.07
   - ***bookstore name*** – *BookMaster*, table *Bookstore*, column *Name*
   - ***sale month*** – *BookMaster*, table *Bill*, column *IssueDate*

7. What are the book sales in relation to the employee's seniority in a given position?
   - **number of copies of books sold** - *BookMaster*, table *Booksale*, column *NumberOfCopies*
   - ***employee's seniority in a given position*** – we define seniority categories as following: *up to a year* (onboarding), *up to 5 years* and *over 5 years* (mature employee). These categories will be calculated based on *CEO EXCEL*, sheet *2*, columns *H* and *I* together with *BookMaster,* table *Bill*, column *IssueDate*. We calculate the position held at the sale moment and the seniority based on difference (in years) between *IssueDate* and date of acceptance for the position (sheet *2*, column *H*)

8. Compare the book sales in relation to the bookstore size, defined as a number employees employed, in current and the previous month.
   - ***number of copies of books sold*** - *BookMaster*, table *Booksale*, column *NumberOfCopies*
   - ***the size of the bookstore*** – we define the categories of the bookstore size as following: *small* – up to 10 employees, *medium* – up to 30 employees and *large* –

over 30 employees. The categories will be determined based on *CEO EXCEL*, sheet *2*, columns *H* and *I* and the *BookMaster*, table *Bill*, column *IssueDate*. We count those employees who were employed at the bookstore at the time of sale.

9. What is the profit in relation to bookstore's distance to shopping mall?[1]
    - ***profit*** – calculated from selling price from *BookMaster*, table *Booksale*, column *Price*. According to the interview, Mr. Book's margin is 7%, so *profit* = *Price*/1.07
    - ***distance to shopping mall*** – there's no such information available in both data sources. The proposals for acquiring such information:
        - enhance CEO Excel with distance to shopping malls
        - distance analysis with the use of tools like zumi.pl (business aggregators)

Analytical problem: "What is the effect of advices given by shop assistants on the sales size?"
1. In how many cases the seller's advice resulted in the purchase of the book?
    - ***purchase of the book*** – *BookMaster,* table *Booksale*, check for entity existence
    - ***whether the seller's advice resulted in the purchase*** – no such information

2. What kinds of books are bought when the seller advises?
    - ***genre*** - *BookMaster,* table *Book*, column *Genre*
    - ***purchase of the book*** – *BookMaster,* table *Booksale*, check for entity existence
    - ***whether the seller advised*** – no such information

3. What titles are promoted by the sellers and what is the profit from these titles?
    - ***profit*** – calculated from selling price from *BookMaster*, table *Booksale*, column *Price*. According to the interview, Mr. Book's margin is 7%, so *profit* = *Price*/1.07
    - ***advised title*** – no such information

It is not possible to build a BI system to support Mr. Book in solving these analytical problems without introducing additional activities in Bookstores. We suggest to introduce a survey system every tenth buyer. Such questionnaire should contain as minimum the following set of questions:
- Did anyone advise the buyer when buying a given book, and if so who?
- Did the seller advise or advise against buying the book you bought?
- Did the seller's opinion affect the decision to buy the book?

This questionnaire is automatically printed when issuing a bill (after giving consent of the buyer) and is completely anonymous. Survey results are automatically uploaded to the survey system and entered into the excel sheet. Sample structure of the survey sheet (each row reflects a single item at the bill):

Column A – BillNumber (text)
Column B – Advising person (empty or name and surname of an employee)
Column C – Advisory (available values: pro, against, no bias)
Column D – Advice influence on buying (numeric from 1 to 5, where 1 – minimal influence, 5 – maximal influence)

---

[1] By blue we indicate the queries that cannot be answered based only on data from BookMaster system and CEO Excel.