Academic Journal of Economic Studies
Vol. 1, No.3, September 2015, pp. 115–125
ISSN 2393-4913, ISSN On-line 2457-5836

Some Considerations Regarding the Design and Implementation of Data Warehouse in Insurance Broker Management

Alexandru Manole

Faculty of Finance and Accounting, Artifex University of Bucharest, Bucharest, Romania
E-mail: alexandru.manole@gmail.com

Abstract *This paper describes a proposal for a data warehouse model, designed for the use in the management of insurance brokerage companies. The model aims to provide information to the leadership of such companies, beyond the classical knowledge drawn from current activity reports. Also, the design process took into consideration the characteristics of the business model analysed. The model is then "exploited" by making some analyses on the data loaded.*

Key words Data warehouse, management, insurance, dimensions, business model

JEL Codes: G22, C88

1. Introduction

In Romania, the insurance broker is a legal entity authorized by Law no. 32/2000, which acts as an intermediary between the insurance companies and the customers, and its business model includes selling insurance policies to customers but also, especially important, providing advisory services for choosing the best solution for protection against insured risks, various issues related to insurance, events that fall under the provisions of the policies, claims related to such events etc.

The business model of the broker is based on its network of customers, so the interest in providing quality and timely expertise for customer needs is mandatory to preserve the image and extend the customer/contracts base.

The insurance brokerage is an important component of Romanian insurance business, a fact that is demonstrated by the financial data, published by various institutions and authorities. According to the annual reports[1] published by the Financial Surveillance Authority (in Romanian, *Autoritatea de Supraveghere Financiara – ASF)*, the insurance brokers have intermediated, in 2013, insurance

[1] Available on http://www.asfromania.ro/publicatii/rapoarte-anuale/rapoarte-asf

premiums totaling 3,928,510,997 lei, while in 2012 the indicator reached a value of 3,603,663,044, which represents a nominal increase of 9.01. In 2011, the total premiums subscribed through the channel of insurance brokers reached 3,115,073,506 lei[2]. In 2014, premiums of 4,553.37 million lei were subscribed through brokers.
The evolution of the indicator "Premiums subscribed through insurance brokers" followed the pattern in figure 1:

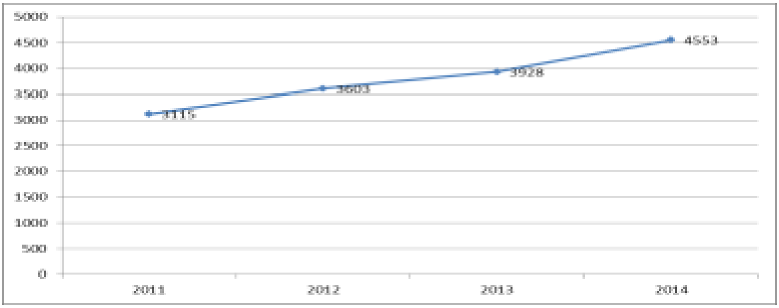

Figure 1. Premiums subscribed through insurance brokers

Note: All values are in million lei

Source: Author's own representation, based on ASF annual reports 2013 and 2014

2. Literature review

The role of brokers in the insurance market is outlined by Rahau *et al.* (2012). Anghelache and Anghel (2014) provide a sound reference for economic and financial modeling. The principles of building and implementing data warehouses were approached, among others, by Popa *et al.* (2006), Manole (2008), Kimball and Ross (2013).

3. Data warehouse design and implementation

Modeling the dimensional data structure for an insurance broker is influenced by the characteristics of its business model. An example of applicable data structure was developed by Popa *et al.* (2006). For our analysis, we shall take into consideration the following elements:

---

[2] Values in this paragraph are drawn from Annex no. I 15 to the Annual Report of ASF in 2013. The report can be found at http://www.asfromania.ro/publicatii/rapoarte-anuale/rapoarte-asf/2147-raport-anual-asf-2013

- The activity measures can be represented by the value of premiums subscribed, and also by incomes achieved by the broker, as commission revenues. These measures are to be included in separate fact tables, because the time reference is not the same. For premiums subscribed, the date at which the policy is issued can be considered, while commission revenues become real measures only after the premium is paid. The premium is assumed and subscribed by the customer when he accepts and signs the insurance policy, the commission revenue is calculated when the customer pays the premium, but is calculated as a percentage of the paid premium, based on a contractual agreement between the broker and the insurance company that issued the policy;
- Time dimension key is to be determined by taking into account the realization of the activity measures. The base level is the date, under the two forms described above. However, there is no need for two separate time dimensions, as all values of the date can be aggregated under a single attribute, the structure of the model allows the correlation with each one of the fact tables;
- The customer dimension is very important, as the client is the key pillar in the business analyzed. This dimension will allow the construction of certain hierarchies, based on the geographical, type (person or legal entity), age group characteristics, gender;
- The agents represent the “sales force” of the insurance broker, their skills are essential not only as sales agents, but as specialists in all topics of insurance, because it can be reasonably expected that clients who need advice will focus at first on the agents;
- The insurance companies, which collaborate with the broker, are also a factor to be considered, as they are part of the contractual and legal base of the business model;
- The types of insurance policies, grouped by categories, this classification must comply with applicable regulations. Also, this grouping supports the periodic reporting activities the broker is mandated to, by law;
- For certain types of policies, the customers can choose to pay the premium in one, or in more rates. Such attribute was included as a separate dimension, to outline the preferences of customers;
- Also, for certain policies, the values are expressed in foreign currencies, and the payment, present or future is expected to reflect this fact (the payment will be made in lei, at an exchange rate that applies in the day of payment). A dimension for currency was included in the model, to reflect the structure of the policies in the portfolio of the broker.

The structure of the conceptual model shall have the form as in figure 2.

The following described aspects are supplemented by the structure of the data source, a relational database designed and tested in a production (running under

MS ACCESS), real environment, whose relational model is compatible with the tables, attributes and roles presented in figure 2. From the source database, the relevant tables were drawn in a data staging area, which is too a relational database, deployed in SQL Server 2014.

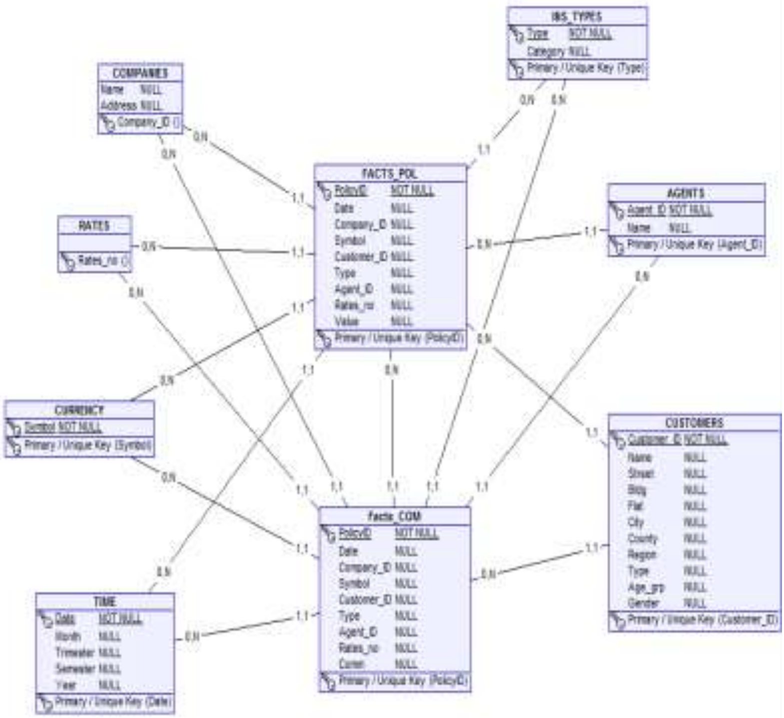

Figure 2. Dimensional model, as designed with Open ModelSphere

The data staging model is presented in figure 3. The model was implemented in SQL Server 2014, with a dataset to be used for test purposes. In order to implement the data warehouse model in SQL Server Database engine, some preparatory operations were made, on the structure presented in figure 3.

The structure of the fact table *FACTS_POL* is based on a table in the data source that includes data on insurance policies. The activity measure, the value of the policy, is drawn as it is from the source database.

The structure of the *FACTS_COM* table was designed in the source database, where the necessary keys were aggregated with the measure (the commission for each rate paid), which was calculated at that step.
As the fact tables were designed and the integrity of data was ensured, two dimensional tables were built based on the facts data: *RATES* and *TIME*. The T-SQL procedures used are presented in figure 4.

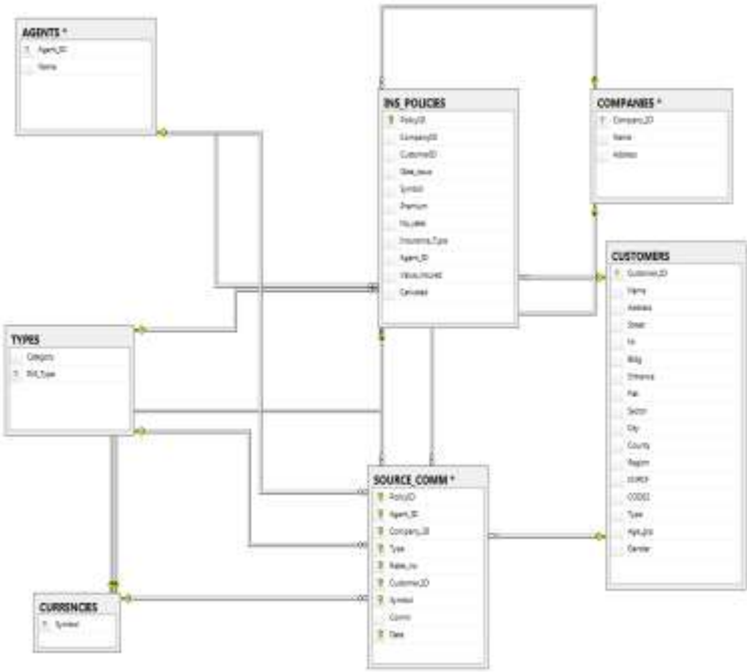

Figure 3. Data staging area, as seen in SQL Server Management Studio

The dimensional table *CUSTOMERS* kept its original attributes, the *Region, Type, Age*_grp and *Gender* were added and loaded with proper values. For example, the gender of persons was obtained based on the Personal Numeric Code (Romanian abbreviation *CNP*), an attribute in the source database, unique for each person and whose initial figure distinguishes men from women. The type of customers was defined by measuring the length of the customer code, of which *CNP* is part, which is fixed (13 figures) for persons and has less characters for legal entities.

```
SELECT DISTINCT FACTS_POL.No_rates AS Rates_no INTO RATES FROM FACTS_POL
SELECT DISTINCT DATE_ISSUE AS Date, MONTH(DATE_ISSUE) AS MONTH,
CASE
    WHEN MONTH(DATE_ISSUE)<=3 THEN 'Quarter 1'
    WHEN  MONTH(DATE_ISSUE)>=4 AND  MONTH(DATE_ISSUE)<7 THEN 'Quarter 2'
    WHEN  MONTH(DATE_ISSUE)>=7 AND  MONTH(DATE_ISSUE)<10 THEN 'Quarter 3'
    ELSE 'Quarter 4'
END
AS TRIMESTER,
CASE
    WHEN MONTH(DATE_ISSUE)<7 THEN 'Semester 1'
    ELSE 'Semester 2'
END
AS SEMESTER,
YEAR(DATE_ISSUE) AS Year
INTO TIME
FROM FACTS_POL

INSERT INTO TIME SELECT DISTINCT DATE, MONTH(DATE) AS MONTH,
CASE
    WHEN MONTH(DATE)<=3 THEN 'Quarter 1'
    WHEN  MONTH(DATE)>=4 AND  MONTH(DATE)<7 THEN 'Quarter 2'
    WHEN  MONTH(DATE)>=7 AND  MONTH(DATE)<10 THEN 'Quarter 3'
    ELSE 'Quarter 4'
END
AS TRIMESTER,
CASE
    WHEN MONTH(DATE)<7 THEN 'Semester 1'
    ELSE 'Semester 2'
END
AS SEMESTER,
YEAR(DATE) AS Year
FROM FACTS_COMM WHERE DATE NOT IN (SELECT DISTINCT DATE FROM FACTS_POL)
```

Figure 4. Generating the structure of the TIME and RATE dimensions, as shown in SQL Server Management Studio

For example, the update query for gender, applied in this case, has the following syntax:

```
Update CUSTOMERS Set Gender='M' Where Left(CNPCF,1)%2<>0 And LEN(cnpcf)=13
Update CUSTOMERS Set Gender='F' Where Left(CNPCF,1)%2=0 And LEN(cnpcf)=13
```

Figure 5. Generating the structure of the level *Gender*

The *CNPCF* attribute is a source column that contains *CNP* for persons and unique identifier code for legal entities. The *CURRENCY* table is drawn from the source, as its records are used for current reporting purpose. The specifics of this table lead to the partial additive character of the *Value* measure. Also, the dimensional tables

AGENTS, COMPANIES and *INS_TYPES* were drawn from the original data source, by selecting only the attributes relevant for the multidimensional model.
Following these operations, the data warehouse model was physically implemented, and is displayed in figure 6.

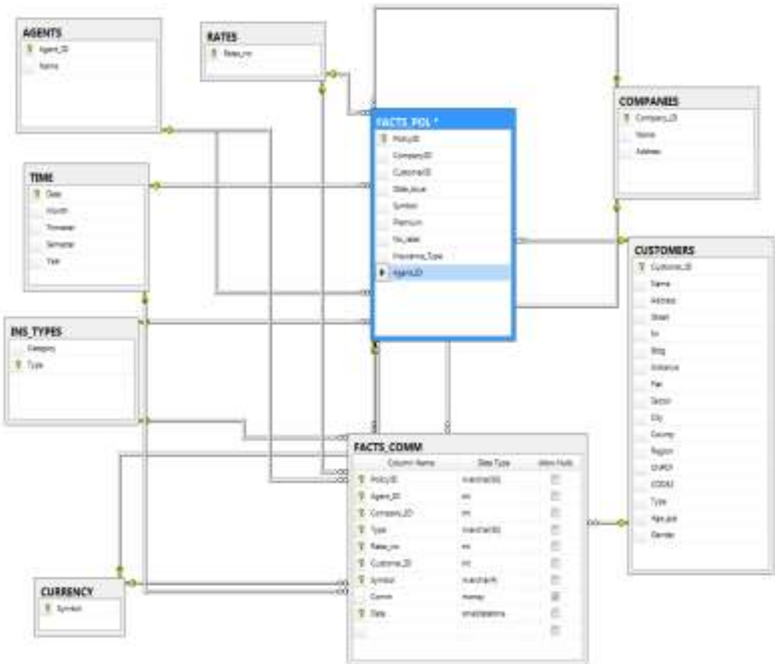

Figure 6. Physical model of the InBroker Data Warehouse, as displayed in SQL Server Management Studio

This model can be exploited with the instruments of SQL Server Analysis Services, allowing the design of cubes and then analyses on data loaded in the data warehouse.

4. Data analysis – additive model

An additive model allows for our data, which can to be summarized across dimensions, to reveal the influence of any modification related to a particular level, or more levels, on the overall modification of the activity measure.

Considering the *Premium* measure, the assertion above might be written under the following formula:

$$\Delta P = \sum_{i=1}^{n} Pi \qquad (1)$$

Where:
ΔPI: variation of the total value of the measure;
Δ Pi: variation of the measure across one dimension level.

This model fits perfectly on the *roll-up* and *drill-down* operations that can be made on the cube. Therefore, multiple analyses are available.
The analysis presented in this section will be focused on the premium insurance measure. However, the same analysis model can be made for other possible measures relevant for insurance business, such as the commission, but also, more useful for big insurance companies, the insured value.
For the purpose of this analysis and on the foundation of the *InBroker* data warehouse, a cube has been designed in Visual Studio data tools, the structure of the cube is outlined in the following diagram:

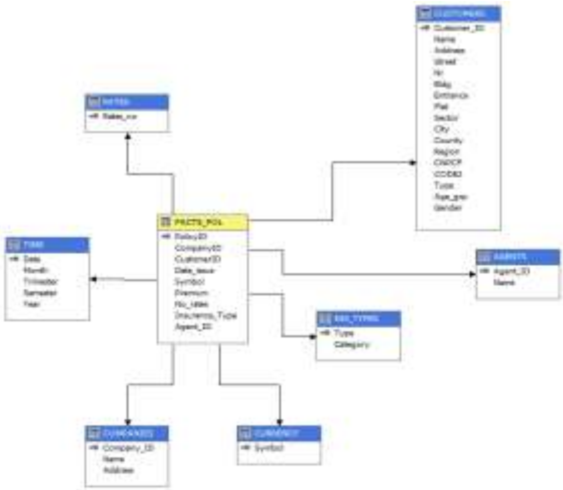

Figure 7. PREMIUM Cube physical model, as extracted from Visual Studio Data Tools

This cube will provide the source data for reports dedicated to premium analyses. The primary step of the analysis is to observe the dynamics of the premium value over time. The report can be represented under the following structure:

Year	Semester	Trimester	Premium	FACTS POL Count
⊟ 2011			1753015.4	6982
	⊟ Semester 1		906338.12	1039
		⊞ Quarter 1	403792.14	504
		⊞ Quarter 2	502545.98	535
	⊞ Semester 2		846677.28	5943
⊟ 2012			1733644.66	12312
	⊞ Semester 1		775488.41	7942
	⊞ Semester 2		958156.25	4370

Figure 8. Evolution of the insurance premiums subscribed (selection from the report) as presented in Visual Studio Data Tools

The report presents both the evolution of insurance premiums and number of policies that were open and paid, totally or partial, during the periods. The report was designed with drill-down capabilities, from the top level (year) to the bottom level included in the multidimensional query, that is *Trimester.*

The analysis could be further expanded to reveal the contributions of agents, who, as specified before, form the interface with the holder of the insurance policy. Such report can have the following form:

	⊞ 2010	⊟ 2011		⊟ 2012	
		Semester 1	Semester 2	Semester 1	Semester 2
1	12493	2252		0	0
10	550	2264	0	2792	144
11	12718.54	19665.48	8092.87	4837	2409
12	36956.024	36042.88	13413.55	327391.13	190424.93
14	7175.69	16286.16	12995.06	10338.06	11835.28
15	34174.45	85414.33	72120.57	66397.38	63617.89
17	7446.58	1328			

Figure 9. Evolution of the insurance premiums subscribed (selection from the report), together with the evaluation of agents' contribution, as presented in Visual Studio Data Tools

The report was based on a MDX query, described below:
*SELECT NON EMPTY { [Measures].[Premium] } ON COLUMNS, NON EMPTY {([AGENTS].[Agent ID].[Agent ID].ALLMEMBERS * [TIME].[Year].[Year].ALLMEMBERS * [TIME].[Semester].[Semester].ALLMEMBERS) } DIMENSION PROPERTIES MEMBER_CAPTION, MEMBER_UNIQUE_NAME ON ROWS FROM [PREMIUM] CELL PROPERTIES VALUE, BACK_COLOR, FORE_COLOR, FORMATTED_VALUE, FORMAT_STRING, FONT_NAME, FONT_SIZE, FONT_FLAGS*

The contribution of agents can be reflected in the number of insurance contracts/policies concluded with the customers, as this indicator reveals the capability of the agents to contact new people and thus enlarge the customer base of the company, and also their ability to promote the individual types of insurance in the portfolio of the company. Thus, the proper report will pursue a three-dimensional analysis: time, type of insurance, agent. The structure of the report proposed is presented in the following figure:

		2010	2011	2012
1		13	2	11
10		1	6	22
11		27	95	47
12	Accidents-persons			
	Auto	69	34	469
	Auto civil responsibility		3200	6528
	Buildings - legal entities	1		
	Buildings - persons	17	9	90
	Cargo			2
	Civil responsibility		1	1
	CMR			1
	Goods - person			2
	Home			5
	Life insurance		1	
	Professional Responsibility			1

Figure 10. Evolution of the number of policies concluded, per agent and insurance type – selection for the contribution of the best agent account (ID 12)

A top level report, without details on agents, can provide the contribution of each type of insurance to the contract portfolio of the company.

5. Conclusions. Future research directions

The proposed model is designed to help the leadership of broker insurance company to take better decisions, based on the data and information the analysis of the "raw material" in the data warehouse can provide. The model is not hard to implement and adapt on the (at least) decisional needs, as each insurance broker must have some type of OLTP application that supports its daily operations, or at least a collection of data files that can be exploited in the process of building such data warehouse. The possibilities of analysis are limited only by the list of dimensions, levels and dimension members, but these elements can be easily customized to accommodate the particular business characteristics of any entity in this field.

The model presented in this paper will form the basis of future research, in which the author intends to present some analyses that can be made on this multidimensional structure, approaching both the analyses on commissions earned by the broker, and also the econometric toolbox in order to draw more useful information for decision makers.

References

Anghelache, C., Anghel, G.M. (2014). *Economic modeling: Concepts, theories and case studies*, Economica Publishing House, Bucharest

Kimball, R., Ross, M. (2013). *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, Third Edition*", John Wiley & Sons, Inc., Indianapolis

Manole, A. (2008). *Software system for modeling financial-accounting decision* (in Romanian language *Sistemul informatic pentru modelarea deciziei financiar-contabile*), Artifex, Bucharest

Popa, Gh. *et al.* (2006). *Microsoft SQL Server*, Economica Publishing House, Bucharest

Rahau *et al.* (2012). Insurance Agents In The Sale Of Insurance - Competence, Responsibility And Limitations Through Legislation, Academica Science Journal - Economica Series, Volume (Year): 1 (2012), Issue (Month): 1 (November), pp. 103-108

msdn.microsoft.com

http://www.asfromania.ro/publicatii/rapoarte-anuale/rapoarte-asf

http://www.asfromania.ro/publicatii/rapoarte-anuale/rapoarte-asf/2147-raport-anual-asf-2013