ISSN 2229-5518

Elegant Solutions
Personal Finance Intelligence

Besim Ismaili

Abstract— Medical doctors always say that the best medicine is the preventive and we as humans work very little on this direction. The subject of this research is not Medicine, but not less important, is Economy.
Financial crisis happened before, is happening now and will happen in future if we do not create a preventive plan. One of the solutions that aim to be part of a "preventive medicine" for financial crises is produced in the data laboratories of ELA (ELegant Analytics) and its name is PFI.
What is PFI?
Personal Finance Intelligence (PFI) represents the name of the Business Intelligence solution for personal finance and planning of your budget. Inspired by the TV Show "Luksusfellen" (A show broadcasted in Norway about people struggling with their economy and living luxurious live-"Luxurious buddy".), this Business Intelligence approach may be a solution for all these who fail to maintain well their own economy, for those who want to perform their economy and last but not least the bank itself.
The purpose of this project is to create a Customer Analytical Cube that would process data for each bank costumer using his/her history for its own benefit and then queries back with the most important answers that customers and the bank itself need.
This solution will include also benchmarking against an Imaginary subject that can be Min, Max or Avg of the customer's values in a set that can be certain region, for a period of time, age group, sex, income ranges etc.
For having more control and planning your own economy, targeting will be an facultative solution when users (bank customers) can put their targets for the costs, a month, quarter or a year ahead and always will be warned when they are about to achieve the amount targeted.
The customer behave will be same important for the bank, so the bank will know what type of customer is and how he/she handle his/her economy.
Project is also meant to be use for the bank itself in cases when they want to evaluate a customer and his/her behave regarding his/her finance stability, because today Credit scoring system lack for some important data that can make decision more accurate. A good example can be credit check, debt history of a customer etc…

Index Terms— Business, Intelligence, Fianance, Crisis, Solution, Modell, Analytics, Data, Customer, Bank, Banks

1 INTRODUCTION

The main focus of this project is the customer, his/her history and his/her behave.
Our first goal is to make possible data collection for customers in the smallest transaction granularity as possible by not impersonating data. This way, bank operates with the whole data diving into details secure and lawfully. Our second goal is by doing Business Intelligence with his historical data to give alerts and advices where he/she is performing bad or giving support where he/she is doing well.
Our third goal is to see where the customer stands, comparing with the region where he/she lives, comparing with his age group, sex and income. This is going to help him improve savings and cut costs by showing how people around him can do with same budget.
Our fourth, but not less important goal is related to the bank itself, where the bank can have clear financial picture for its customer and can decide much better than credit scoring system.
This system lowers the risk and improves the loyalty with customers.

Project Highlights:

Data center that has the capacity to

- Data transfer once a day or live-data (for the bank side)
- Centralized Customer Intelligence for the Entire Bank
- Live Data transfer and access
- Separate service for each client
- Client vs. Average, Max or Min of a set of clients (Benchmarking)
- Other Intelligence analysis (Geography, age, sex etc…)

Processes on the fly (Administration and Maintenance):

- Optimizing ETL
- Optimizing DB and DWH
- Optimizing Indexing and Data Volume
- Query Performance regarding MDX calculated measures

It is very recommended that privacy and security issues regarding credential information about customers to be distributed in a high consideration.

Also users' impersonation with data source and data reported is highly recommended to be solved in the best possible way, including data source security till Cube role group's security.

Other recommendation is regarding planning the data volume and performance upon queries requests in a large volume of data in production environment.

Data Architecture of the PFI Solution

Database is the main source of data and is usually an accumulator of data entry and daily transactions (manually or automatized). In our case, this Customer DB is the main data source for bank costumer data and for the solution itself. Usually is made by tables that have different information and transaction data. The problem with it is that DB has raw data and is not organized, calibrated and optimized for the need of business. It is normal that we need a sophisticated ETL process that will make this raw data usable.

So from the database we create or transform a Data warehouse.

Data Warehouse (DWH) is a database too, but its structure is enough organized to answer business requests efficient in time and respecting data quality. Therefore, we must transform the raw data Database to DWH and this can be done in many ways using Store procedures, SSIS as a tool or other tools from different platforms. Since I am a Microsoft BI Expert, I would prefer using Microsoft SQL BIDS SSIS to do these transformations for many reasons that could be price, software/hardware compatibility and user experience factors.

The structure of a DWH is very important to a successful BI solution, while is not that important how raw data in DB is organized. DWH is organized around fact tables and dimensions, that's why we referrer as Dimensional Modell. In fact you find numeric data such amount, quantity and foreign keys to dimensional data such FK_CustomerID, FK_Date, FK_Product, FK_BankBranch, FK_BankSeller etc…

Based on previous experiences, complex data structures are usually organized in Snowflake Schemas, would be enough for data to be presented in right way and it would approximately look like fig 1.1.

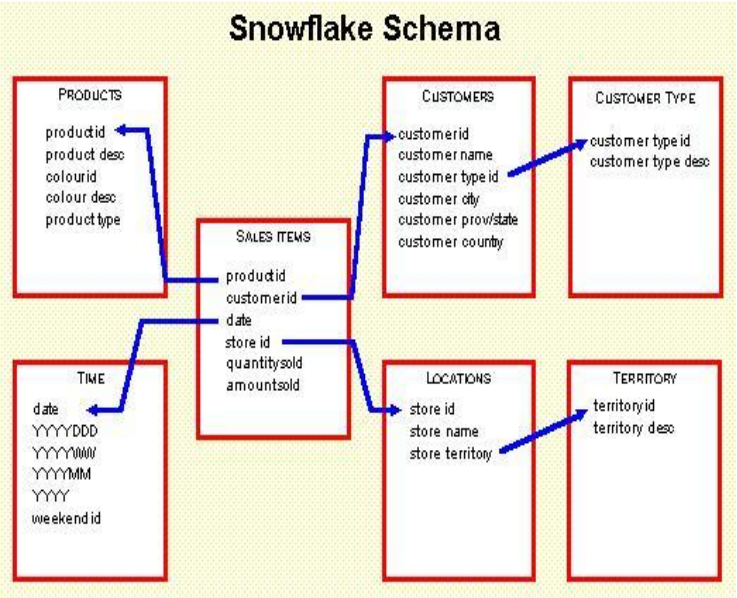
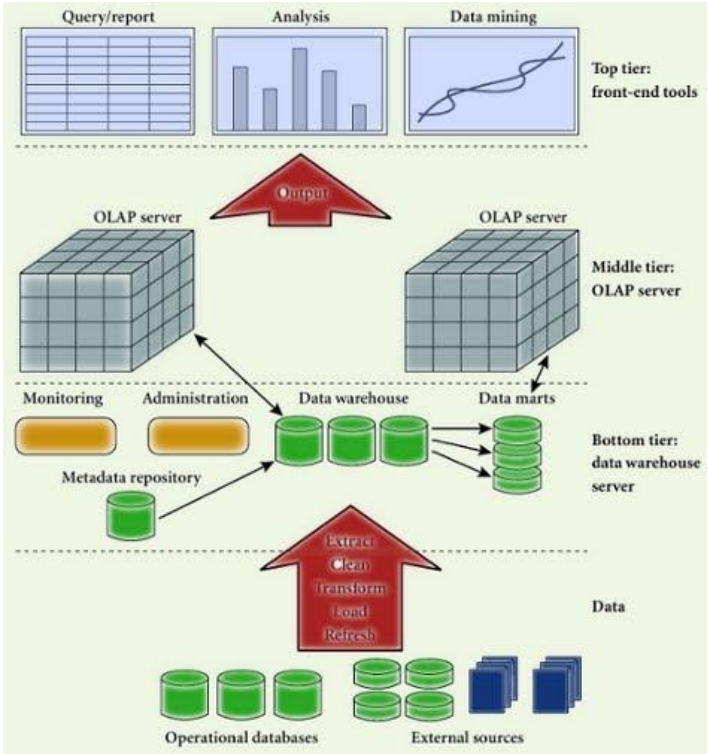


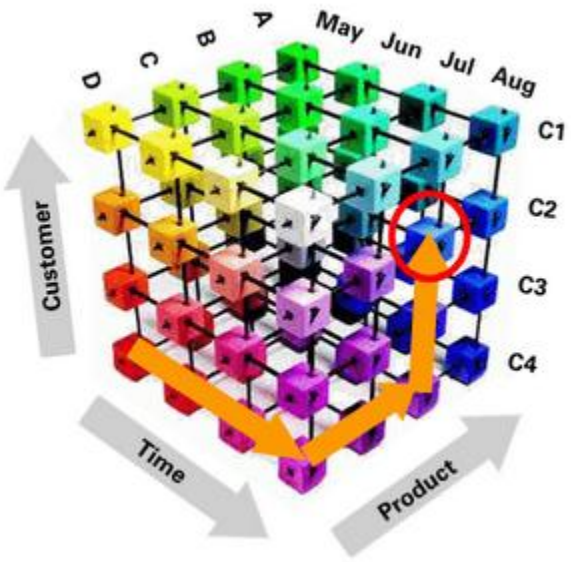
FIG 1.1. Snowflake Schema of a simple Costumer Diagram

This is the core diagram of the solution and represents the most important structure where the technicaly gives shape to the whole project.

With this structure in the heart, the general solution can be visualized as:



2. Cube Structure (OLAP) and MDX for the Solution



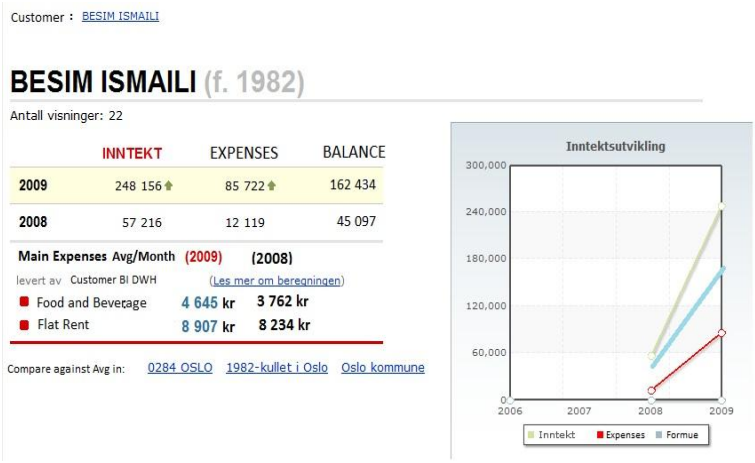
Cube structure will not be that complex as it looks in first time. After defining the business logic and needs then constructing the OLAP Cube is very easy.

The structure of every OLAP Cube is based on Dimensions and Facts.
The common dimension of all Cubes is Time Dimension, often has members such Year, Quarter, Month and Date and for special case its granularity can drill to the transaction level. Because, we deal with daily based transactions our granularity would be Date (in some cases my drill deeper to hours).
We are going to have a main dimension Customer which will have Attributes such Customer Name, Customer ID (Unique and Key), Post Code, Address, City, Area, Region, Country, Subscription Date (Date joined the Bank), Flag for Active or Not active Costumer, and will be handled Slowly Changing Dimension type 2- SCD2.
A Geography dimension would be necessary for filtering location data such Country, City, Region, etc… Account dimension is very important because it will note the transaction to which account it belongs, Account Dimension may have data such Account Number and Account parent indicating which accounts for income as parent and then child salary, renting apartment, income from investment, shares and then parent Expenses with children Food and beverage, Electricity, house rent, mortgage, loan etc…
To make clear what the Cube will do in background when I query is sent from a user, we are going to have a sample query in MDX as follow:
If you want to see how much he spent in Food and Beverage for whole Mai 2011 and compare it with same month last year (Mai 2010)? Here is the description for this problem:

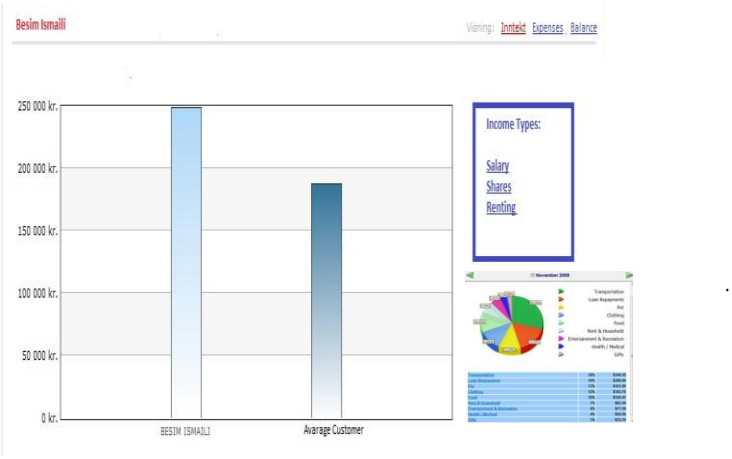
UserID as unique with value for each customer accessing the cube: [Costumer].[UserID].CurrentMember) can see [Measures].[Food and Beverage]=([Measures].[Amount],[Accounts].[Expenses].[Food and Beverage]). ([Measures].[Food and Beverage], [Time].[Month].&[Mai2011]) for May 2011 and ([Measures].[Food and Beverage], [Time].[Month].&[Mai2010]) for May 2010, related to his own user and against Ola Nordman (Avg user) for Benchmarking.
The difference is obvious: First Query - Second Query and you can get easy in Percent as well: (First Query - Second Query)/ First Query in "%" format.
But how we get the Average Customer (Ola Nordman)?
For Example: Average Customer of the City Oslo means, Sum([Measure].[Food and Beverage], [Geography].[City].&[Oslo])/Count([Customer].[UserID].members, [Geography].[City].&[Oslo). So we have the average expenses of Food and Beverage in Area Oslo and every user from Oslo can do benchmarking against the model provided.

3. GUI Screen Shots of the Front End Solution

This is just an imaginary case of a customer (this time me) showing how the GUI may look like.
The GUI should be as much as possible User Friendly and has the most important filters visible and easy to find.
This can be built in any web portal solutions of e-banks as a simple web part where the Source is in the OLAP structure discussed above.
The functionality and design of GUI can be based on any .Net Framework programming language.



Here you can see a simple benchmarkbeetwen a Customer Bank (me) and Avg subject that represent the Avarage Set of my City.



CONCLUSION

If the research papers will have the magnitude to reach enough audience to support this idea, it will be great to implemment in any bank at any plce in the World. I think that many banks will use many tactics against this solution which give power to their customers, enough to organize and maintain their economy. Also, they may find legal issues for using too much data and inpersonalization cases.
As you can see, this solution to be implemented has not just technical issues, but political issues as well and we need to be prepared for a political battle.

REFERENCES

[1] Figures are *taken from Internet and used only for illustration, www.google.com*
[2] The text is full originally written and there is no external reference.