

Using Business Intelligence Tools for Predictive Analytics in Healthcare System

Mihaela-Laura IVAN

Department of Economic Informatics and Cybernetics
Bucharest University of Economic Studies
Bucharest, Romania

Manole VELICANU

Department of Economic Informatics and Cybernetics
Bucharest University of Economic Studies
Bucharest, Romania

Mircea Răducu TRIFU

Department of Economic Informatics and Cybernetics
Bucharest University of Economic Studies
Bucharest, Romania

Cristian CIUREA

Department of Economic Informatics and Cybernetics
Bucharest University of Economic Studies
Bucharest, Romania

Abstract—The scope of this article is to highlight how healthcare analytics can be improved using Business Intelligence tools. Healthcare system has learned from the previous lessons the necessity of using healthcare analytics for improving patient care, hospital administration, population growth and many others aspects. Business Intelligence solutions applied for the current analysis demonstrate the benefits brought by the new tools, such as SAP HANA, SAP Lumira, and SAP Predictive Analytics. In detailed is analyzed the birth rate with the contribution of different factors to the world.

Keywords—Healthcare Analytics; Business Intelligence tools; SAP HANA; SAP Lumira; SAP Predictive Analytics; Birth Rate; Big Data

I. INTRODUCTION

According to a 2016 survey by the World Economic Forum [1], today's society is experiencing the fourth industrial revolution. A phenomenon that led to significant changes in the labor market by creating new jobs, while in a few years ago it was not possible or were unimaginable. The evolution of new technologies such as Business Intelligence, Big Data, Cloud Computing, Mobile Programming, Social Networks, Cyber-security and others have radically changed the way people communicate and collaborate. Educational programs offered by academic institutions must adapt to this technological revolution so that graduates are prepared for the new requirements imposed by the labor market, economic and social change. Areas such as healthcare, accounting, marketing, management, tourism, and many others have another perspective in this context because all were or are being influenced by IT&C technologies. Online platforms have emerged by which people perform medical appointments, electronic payments, reserves tickets, holidays and vacations reserves and many others. These operations in the past were taking place in bank branches or travel agencies.

Business Intelligence is very timely, being promoted both by large technology companies producing technology and professional associations fostering scientific research, such as ACM (Association for Computing Machinery) and IEEE (Institute of Electrical and Electronics Engineers). On 16

March 2016, ACM organized a Webinar on "From BI to Big Data - Architecture, Ethics and Economics", which shows that the domain is very relevant at this time.

The creative use of Big Data, especially for Internet of Things (IoT) transforms business models by supporting start-ups and strengthening or destruction of existing business values. Many supporters of technologies referred focus to use Big Data in the marketing area, but the greatest value will come from daily using and from new ways of transformation. To make appropriate technological and architectural choices is vital in designing and managing automated environments that Big Data should have and that the Internet of Things will ask. However, the real deal requires a close look at the ethical and economic aspects that Big Data poses. They concern personal privacy, employment and social disruption which must all be addressed urgently if the individual businesses and society support successful navigation in this transformation based on data for all aspects of business and technology [2] [3].

II. BUSINESS INTELLIGENCE TOOLS FOR ANALYTICS

The technologies analyzed are implemented in SAP HANA and above this SAP HANA In-Memory database is deployed SAP Lumira or SAP Predictive Analytics. For this research, SAP Lumira [4] is used for illustration of factors affecting the birth rate in a country and SAP Predictive Analytics to see the contributions of the analyzed inputs into birth rate output with classification model, where the factors are classified depending on significance and contribution. The heart of HANA application development contains data and views. The main feature of SAP HANA is the ability to run different information views. The three possible information views that can be designed with the help of SAP HANA are presented in Table 1 [5].

TABLE I. COMPARATIVE ANALYSIS ATTRIBUTE VIEWS/ANALYTIC VIEWS/CALCULATION VIEWS

Attribute Views	Analytic Views	Calculation Views
- Used to model an entity that is based on joins between various	- Many have two types of columns: attributes and measures;	- Calculation Views provide the combination of tables

<p>source tables and then to choose for output the relevant table's columns and rows;</p> <ul style="list-style-type: none"> - Could model descriptive attribute data by using only attributes; - This type of view use the join Engine; 	<ul style="list-style-type: none"> - Can contain attribute views inside of them to achieve more depth of attribute data; - Use the OLAP Engine in general, but if the view contains calculated attributes then it is used the Calculation Engine; 	<p>with other views like attribute views and analytic views;</p> <ul style="list-style-type: none"> - The nodes that can be created with calculation views are union, join, projection, aggregation, and rank; - Two types of calculation view can be created: script calculation view using SQLScript and graphical calculation view;
--	---	--

Data sources for SAP Lumira can be Excel, SAP HANA Analytical Views, Software Development Kits (SDK) and many other data sources. SAP Lumira can connect data no matter where the data lives.

SAP Lumira's strength is the possibility to visualize data in different friendly ways like interactive maps, beautiful charts, and infographics. With the help of SAP Lumira, BI analysis with intuitive dashboards, and securely share insights and data stories can be shared with decision makers using SAP Lumira server and cloud platforms with browser and mobile-based experiences to further analyze data and collaborate with colleagues on datasets, stories, and other business intelligence artifacts.

One of the SAP HANA analytical views used to highlight the population analytics is presented below in Figure 1 [6].

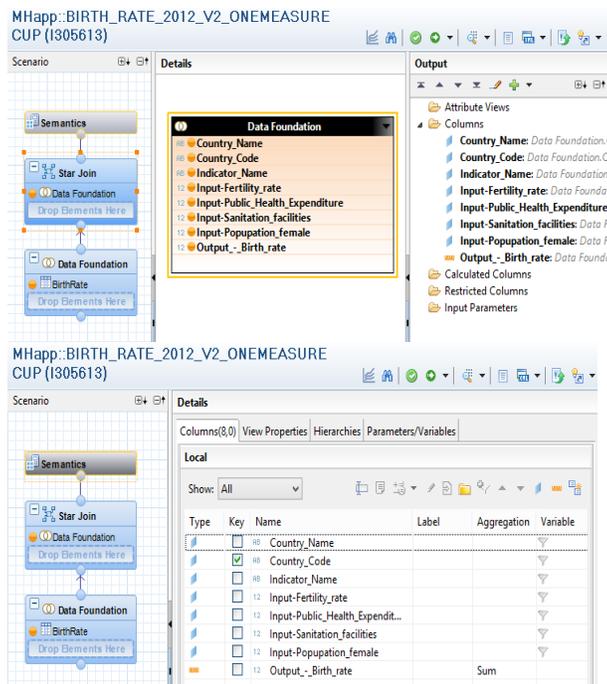


Fig. 1. SAP HANA Analytical view (used in SAP Lumira)

As can be seen, the table BirthRate was added into Data Foundation and all the columns were added for the output.

After that a new dataset is created in Lumira to connect to the presented analytical view. SAP Lumira is online connected to SAP HANA and the download of data from SAP HANA is made offline. The dataset have the columns from analytical view's output separated in measures and dimensions as can be shown in Figure 2.

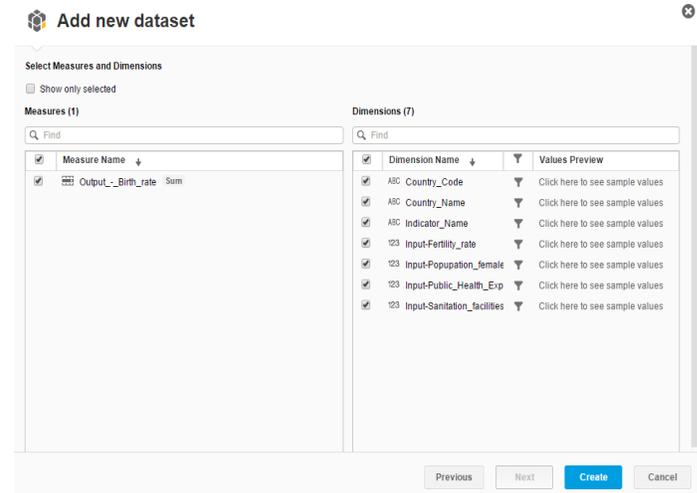


Fig. 2. SAP Lumira dataset of a SAP HANA analytical view

SAP Lumira datasets can be shared, exported or printed to collaborate with colleagues on the same datasets. The dataset sharing methods are the following:

- Save the dataset as .csv file or xls file;
- Publish the dataset to SAP HANA to create a new analytic view;
- Publish to SAP BusinessObjects Explorer. In this way the dataset can be used as an Information Space data source;
- Publish a dataset to SAP Lumira Cloud, to save documents and work together with colleagues on datasets;
- Publish a dataset to SAP BusinessObjects Business Intelligence platform;

III. BIRTH RATE ANALYTICS WITH SAP HANA AND SAP LUMIRA

The definition of Birth Rate indicator is the number average number of births for every 1000 people in a country. The birth rate indicator, BR, is calculated based on the formula below:

$$BR = \frac{NLB}{TP} \times 1000 \quad (1)$$

where:

NLB – the number of live births

TP – the total of population

The analysis is of 207 countries with data related to birth rates for 2012 year.

The influences factors for birth rates presented in this paper are the following [7]:

- Fertility rate – is the number of children that could be born by a woman if she were to live to the end of her childbearing years and bear children in correspondence with her current age-specific fertility rates;
- Public health expenditure – is the recurrent and capital spending from government budgets, different borrowings and grants and social health insurance funds;
- Sanitation facilities – is the possibility to access the improved sanitation facilities. It is referred as the percentage of the population using improved sanitation facilities;
- Female population - is the proportion of the population that is woman;

In reality, there are more external factors affecting population depending on each country and region like the following factors, which will be presented on future researches [8]:

- Capital income - is income that comes from capital for a country;
- Age-sex structure - the distribution of the population by sex and age;
- Religious beliefs and social beliefs regarding contraception and abortion;
- Economic prosperity - in terms of the correlation between the economy growth and the willing of families to have more children;
- Female employment - Employment rate of women referred to percentage of female population;
- Poverty levels - babies are seen as a method for developing countries because they will become to earn money;
- Infant Mortality Rate - depending on the country's IMR, if it is high then some of the children will die;
- Typical age of marriage- is the age at first marriage [9] [10];

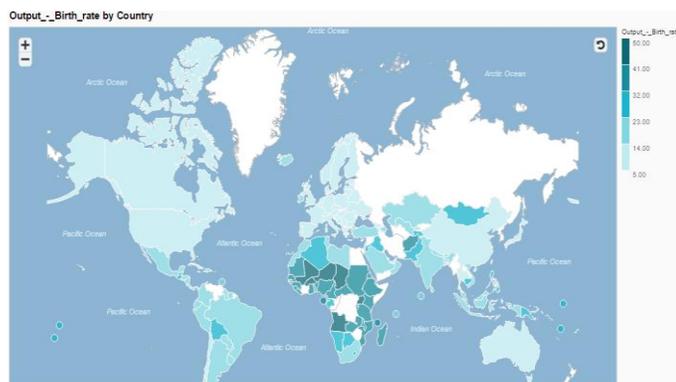


Fig. 3. Birth rate by country

The above image, Figure 3, shows the birth rate by country for 2012 year. This map was created with SAP Lumira where the global natality status is emphasized by color. As a first impression can be seen that Africa has the biggest birth rate, with the following three top countries: Niger (49.87), Angola (46.50), Mali (44.75). While the countries with the lowest level of births are Japan (8.20), Germany (8.40) and Portugal (8.50). Romania is part of the countries with low level of birth rate, the value is 8.80, which almost as small as the first range value.

In the following section are analyzed how the selected inputs affect the number of births. Figure 4 and Figure 5 present two tables for comparative analysis on how fertility rate indicator, female population indicator, public health expenditure indicator, sanitation facilities indicator make birth rate to increase or to decrease. One table contains the countries having the biggest birth rates and another one the countries having lowest birth rates. These two tables were obtained with SAP Lumira's ranking functions, top three and bottom three on a crosstab control as displayed in Figure 4 and 5. A crosstab allows complex multidimensional analysis which is useful to view the exact values or to examine data from multiple measures.

Top 3 Output_-_Birth_rate by Country_Name, Input - Fertility Rate, Input-Population_female, Input-Public_Health_Expenditure, Input-Sanitation_facilities

Country_Name	Input - Fertility Rate	Input-Population_female	Input-Public_Health_Expenditure	Input-Sanitation_facilities	Output_-_Birth_rate
Niger	7.64	49.63	33.06	10.10	49.87
Angola	6.25	50.42	62.16	49.10	46.50
Mali	6.40	49.56	38.83	23.30	44.75

Fig. 4. Top three Birth rate by fertility rate, population female, public health expenditure and sanitation facilities

Let us take a look on how the number of births occurring during one year, per 1000 population estimated at mid-year is correlated with the inputs for Niger country. We want to highlight the factors with the biggest influence and the reasons for this big birth rate.

Fertility rate that is calculated as $\frac{\text{number of babies}}{\text{woman}}$ has the greatest value of 7.64. For this country fertility rate is strongly correlated with birth rate.

Female population - proportion of population that is female, with 49.63% value. This proportion shows that the population is uniform divided and this argues the high value of birth rate.

Public health expenditure - percentage of total health costs is 33.06%, which means that the total expenditure on health by Niger is not so high and surprising this factor does not affect the number of births.

Sanitation facilities - proportion of population with access to sanitation, to safe water and those who practice good hygiene represent only 10.10% [11].

As a conclusion of the above analysis, the output value (49.87) of birth rate is strong influenced by fertility rate, female population. While the two other inputs, public health

expenditure and sanitation facilities, does not affect the number of births. Although, the factors like religion and poverty are also important factors for birth rate, but these will be considered in future researches.

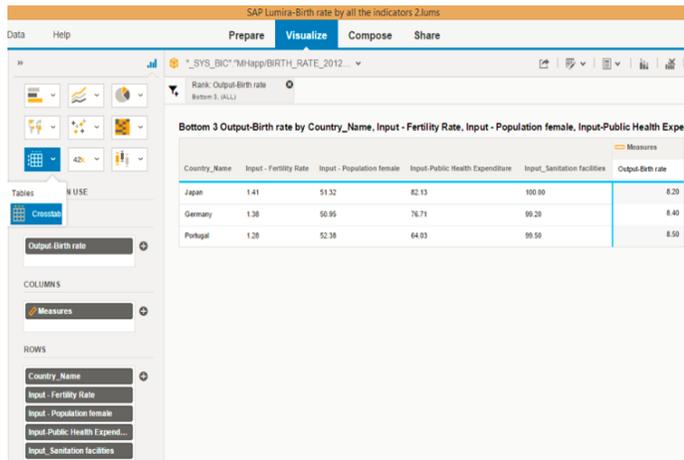


Fig. 5. Bottom three Birth rate by fertility rate, population female, public health expenditure and sanitation facilities

As a comparison to Niger, Japan has the littlest birth rate value, 8.20. The input factors for this output result are the following:

Fertility rate, the average number of children that a woman may give birth in her lifetime, with value 1.41 is also strongly correlated with the low value of birth rate.

Population female is 51.32 %, which is inversely correlated with birth rate and is not so relevant for the output.

Public health expenditure with the rate of 82.13%, which is a high value. Even if a big part of government budget is granted for the health sector, this endeavor is not enough to help the number of births to growth.

Sanitation facilities with the rate of 100.00% emphasize that all the population has gained access to improved sanitation facilities.

Further, in Figure 6, it is calculated with classification model of SAP Predictive Analytics the variables contribution into birth rate output.

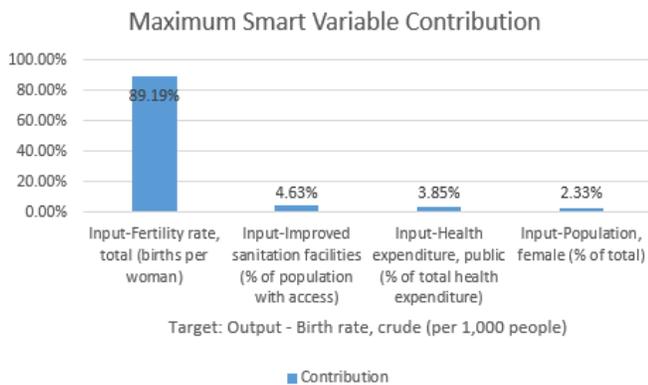


Fig. 6. Inputs contribution into birth rate output

This chart confirms the above assumptions. Also, the figures demonstrate that government efforts to help the low birth rate have not growth the desired effects. In reality, the government intention will not bring immediate fruit to straighten the nations demographic. That is why policymakers have to sustain the needed steps without being affected by down figures.

The trust of the model is calculated with Predictive Power indicator, abbreviated as KI, and Prediction Confidence indicator, abbreviated as KR. Prediction Confidence refers to the ability of the model to achieve the same performance on a new set of data. Models are power if the value is ≥ 0.95 . While, Predictive Power, describes the percentage of information which exists in the target variable. Is good when the Predictive Power value is higher. That model with the highest sum of Predictive Power and Prediction Confidence will be chosen because of the robustness, accuracy and coherency it has. The range of possible values for these two predictions indicators is between 0 and 1. KI and KR values of classification model can be seen in Figure 7.

Model Overview

Overview

Model: Output - Birth rate, crude (per 1,000 people) - "BirthRate2013"	
Data Set:	I305613."BirthRate2013"
Initial Number of Variables:	9
Number of Selected Variables:	4
Number of Records:	207
Building Date:	2016-04-23 12:43:20
Learning Time:	0 s
Engine Name:	Kxen.RobustRegression
Author:	I305613

Continuous Targets (Number)

Output - Birth rate, crude (per 1,000 people)	
Min	7.9
Max	45.745
Mean	21.829
Standard Deviation	10.178

Performance Indicators

Target: Output - Birth rate, crude (per 1,000 people)

rr_Output - Birth rate, crude (per 1,000 people)	
Predictive Power (KI)	0.9749
Prediction Confidence (KR)	0.9813

Variable	Maximum Contribution	KI	KR	KI + KR
Input-Fertility rate, total (births per woman)	89.19%	0.982	0.988	1.97
Input-Improved sanitation facilities (% of population with access)	4.63%	0.7946	0.9505	1.7451
Input-Health expenditure, public (% of total health expenditure)	3.85%	0.4655	0.934	1.3995
Input-Population, female (% of total)	2.33%	0.4168	0.9312	1.348

Fig. 7. Performance indicators

Also, valuable information from the model overview is regarding Min, Max, Mean, and Standard Deviation which extract relevant targets from the data set, very useful for analytics and decision makers.

IV. CONCLUSIONS

With the help of the new tools like SAP HANA, SAP Lumira and SAP Predictive Analytics data can be easily aggregated to make comparative or predictive analysis useful for important domains like it is healthcare field. In other words SAP Lumira 'predict the unpredictable' with the capability to load data from all sources, present valid insights and as a result to help on better decisions.

The results presented show the existence of the relationship between fertility rate in principal, female population, public health expenditure, sanitation facilities and birth rate. This relationship between four inputs and one output behaves differently among the countries due to geographical aspects, labor market characteristics and their welfare state model.

From our point of view, the humanity is in a change process. The lifestyle of people is tending to be robotized, even the process of making babies can in our time be helped mechanically with 'In Vitro Fertilization' process which brought five million children in 35 years [12].

REFERENCES

- [1] World Economic Forum, The Future of Jobs: Employment, Skills and Workforce Strategy for the Fourth Industrial Revolution, January 2016, Available at: http://www3.weforum.org/docs/WEF_FOJ_Executive_Summary_Jobs.pdf
- [2] B. Nedelcu, "Business Intelligence Systems", *Database Systems Journal*, vol. IV, no. 4/2013, pp 12-20.
- [3] V. Farrokhi and L. Pokorádi, "The necessities for building a model to evaluate Business Intelligence projects- Literature Review", *International Journal of Computer Science & Engineering Survey (IJCSES)*, Vol.3, No.2, April 2012, pp. 1-10.
- [4] D. Crockett, "Defining Predictive Analytics in Healthcare" [Online], 2013, Available: <https://www.healthcatalyst.com/predictive-analytics>
- [5] ***, SAP SE or an SAP affiliate company, SAP HANA Modeling Guide, 2014.
- [6] M. L. Ivan, "Improving Business Intelligence Applications by Using New Generation of Web and Mobile Technologies," *Informatica Economică Journal*, Vol. 19, No. 4, 2015, pp. 81-89.
- [7] ***, World Bank [Online], Available: <http://data.worldbank.org>
- [8] E. Jaba, I. A. Chirianu, C. B. Balan, I. B. Robu, M. D. Roman, "The Analysis of the Effect of Women's Participation in the Labor Market on Fertility in European Union Countries Using Welfare State Models", *Journal of Economic Computation and Economic Cybernetics Studies and Research*, No. 1 2016, Vol 50, pp. 69-84.
- [9] ***, PAPP – Glossary, http://papp.iussp.org/sessions/glossary/glossary.html#population_pyramids
- [10] ***, OECD library, http://www.oecd-ilibrary.org/employment/employment-rate-of-women_20752342-table5
- [11] ***, International decade for action 'Water for Life' 2005-2015 <http://www.un.org/waterforlifedecade/sanitation.shtml>
- [12] ***, The amazing story of IVF: 35 years and five million babies later [Online], Available at: <http://www.theguardian.com/society/2013/jul/12/story-ivf-five-million-babies>.