An Examination of Factors Influencing the Quality of Data in a Data Warehouse

Nouha ZELLAL and Abdellah ZAOUIA,

Doctoral Studies Center, National Institute of Posts and Telecommunications, Rabat. Morocco

Summary

Data quality in a data warehouse is a key success factor for each Business Intelligence project. In fact, it has a direct impact on taken decisions. If the data quality is good enough for decision makers, the decision support system is very helpful for them. It allows them to have the right inputs to take the right decisions wherever and whenever they need them. But when the data warehouse is of poor data quality, it can have serious impacts on taken decisions that may be even disastrous.

Considering this importance of data quality in data warehouse, we aim in this study to investigate the influence of such contingency factors as top management commitment, data quality management practices, external expertise, data quality at the source, teamwork and technology factor, on the one hand, and data quality in data warehouse, on the other.

We developed a conceptual model where we formulated the relevant hypotheses (Zellal & Zaouia, 2015) and then we established the measurement model (Zellal & Zaouia, 2016). We conducted the survey in Morocco and we used a structural equation modeling technique to analyze the collected data.

The objective of identifying the most critical factors is to enable stakeholders to better use their scare resources while implementing a data warehouse by focusing on these key areas that are most likely to have a greater impact on the data quality in data Warehouse.

Key words:

Data Quality, Business Intelligence, Data Warehouse, Influencing factors.

1. Introduction

Since data warehouse has been coined by Inmon in 1990, it has known a great expansion in the IT world. It is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process (Inmon, 1992). It allows organizations to consolidate and summarize data from different systems in order to have only one decision support system (DSS).

The data available in the DSS should be accurate enough, timely enough and consistent enough for the organization to survive and make reasonable decisions (Orr, 1998) .In other words, the data in data warehouse should be of good quality to support the decision making process. In this context, we define the data quality in data warehouse as defined in the data quality model standard ISO/IEC 25012: "The quality of a data product may be understood as the degree to which data satisfy the requirements defined by the product-owner organization".

In the case of poor data quality in data warehouse, the managers may take the wrong decisions. Hence, decision support system may have adverse consequences and impact negatively on the performance and the benefits.

That's why poor data quality has been considered, both in literature and by practitioners, as one of the factors that cause the failure of data warehouse (Briggs, 2002).

Because of the huge importance of data quality in data warehouse and because of the lack of academic research concerning data warehouse success, there has been a call for rigorous empirical studies to examine data warehousing success factors (Lee, Lee, & and Suh, 2001).

In this context, the purpose of this study is to examine the factors influencing the data quality in data warehouse. We begin in the first section by presenting the hypotheses of our reviewed research model. In the next section, we present the adopted research methodology and the data analysis. Finally we discuss the results.

This study is interesting on two levels:

- On theoretical level, this research will highlight the influence of different contingent factors on data quality in data warehouse.
- On a practical level, it will help the practitioners and the stakeholders to focus on the most significant factors influencing data quality in data warehouse in order to build a decision support system of high data quality.

2. Literature background and Research Model

Despite the recognition of data warehouse as a strategic information source for decision makers, academic research has been lacking concerning data warehousing practices and its critical success factors (Shin, 2003).

Hence, to find the factors influencing data quality in data warehouse, we referred first to different articles concerning data warehousing, data quality management and information systems implementation. Then, the research model has been reviewed and updated after discussions with researchers in the same field (while international

Manuscript received August 5, 2017
Manuscript revised August 20, 2017

conferences WCCS 15 and CIST 16) and after assessment of factors influencing data quality in data warehouse in 3 large firms in Morocco: ONCF (Railways National Office), ANP (National ports Agency) and OCP (National Office of Phosphate).
In this paper, we prensent the reviewed model that we examined empiracally. The research model can be devided to structural model and measurement model:

- The structural model relates the latent variables: Data quality in data warehouse and the influencing factors
- The measurement model relates the measured variables to latent variables

2.1 Structural research model

In this section, we present the hypothesis built in our research model as shown in the figure below.

Figure 1: Structural research model

2.1.1Data quality in source systems

The source systems are the inputs of the data warehouse system. This last one is just a logically and physically transformation of multiple operational source applications. That's why the data quality in data warehouse is dependent of the data quality of its input data even if this influence can be moderated by the data warehousing process. For example, in the case of multiple data sources, data quality is impacted by the synergy between the different sources. The data quality issues may be either on the schema level when data models and schema designs are heterogeneous, or on instance level such as semantic heterogeneity (Amit & Emilie, 1999) or varying timeliness of data sources.
Furthermore, the quality of data particularly in the source systems was considered in literature review as crucial for any Business Intelligence system implementation because of its impact on the quality data available in the data warehouse (Yeoh & Koronios, 2010).
Thus, it is hypothesized that:
H1: Data quality in source system (s) influences the data quality in data warehouse.

2,1,2 Data quality management practices

As defined by Weber et al. data quality management is the quality-oriented management of data as an asset, that is, the planning, provisioning, organization, usage, and disposal of data that supports both decision-making and operational business processes, as well as the design of the appropriate context, with the aim to improve data quality on a sustained basis (WEBER, OTTO, & OSTERLE, 2009)
Organizations that are adopting data quality management practices are referring to practitioner's guides to analyze their data, to analyze the data quality requirements, to identify the critical areas of data and to evaluate the cost of data quality. They assign data responsibilities, assess data quality, improve it and monitor it.
So in the light of this, we hypothesize that:
H2: The adoption of data quality management practices improves the data quality in data warehouse.

2.1.3 Top Management Commitment

Today, most companies delegate authority for managing data quality to the IT department and Data warehousing Team. Although IT must be involved in the process, it doesn't have the clout to change business processes or behavior that can substantially improve data quality (Orr, 1998). It is up to top management to set up data quality goals according to decision makers' needs and task decisions. It is top management duty to set up policies for data quality and to allocate resources to achieve the data quality goals.
Reviewing the literature, we found that the strong commitment of top management is a key factor to have a high quality product, whatever its type (Dale & Duncalf, 1985). This link between top management commitment and quality of product should also apply in the context of data warehouse which is an Information Product (Shankaranarayanan & Cai, 2006).
Focusing on the implementation of information systems, Watson and others proved that top management support is critical to all major IS initiatives and noted its importance in data warehouse development as well (Haley, Watson, & Barbara, 1997) (Wixom & Watson, 2001)
On the light of these works, we propose:
H3: The higher the top management supports data warehouse implementation, the greater is data quality in data warehouse.

2.1.4 External Expertise Quality

We mean by External expertise the external mediator's entities such as the BI vendors and IT consultants, who take in charge the development of the target solution, provide the training, maintenance and technical support for companies implementing data warehouse.

Surely, developing a data warehouse requires skills and deep knowledge. It requires both technical and business expertise. That's why, the external experts must give the best of their knowledge, experiences and competencies in order to build a data warehouse of high data quality.

Understanding very well the business requirements, taking into account the development environment and referring to their knowledge, external experts should advise the company implementing the data warehouse to use the adequate data profiling, data quality and ETL tools. They can implement validation routines, data quality checks and metadata repository.

In this context, Ifinedo confirms that the quality of external expertise influences the quality of the information generated by the information system (Ifinedo, 2008) .And since the data quality in data warehouse is dependent to data quality at the source, the external experts can only moderate this impact.

That's why the company implementing data warehouse should give the biggest importance to expertise quality while selecting the vendors and consultants, and not to refer only to the price and time of data warehouse implementation.

In addition to this, Alhyasat considers vendors and consultants as a support quality factor in his Data Warehouse Success Framework (Alhyasat, 2013) .And Thong et al. found that vendor support and consultant effectiveness are closely related to the overall information system effectiveness (Thong, Yap, & Raman, 1996) .Wang et al. also consider the system provider as an important factor in the establishment and maintenance of a quality system (Wang, Shih, Jiang, & Klein, 1996).

On the light of these works, we propose:

H4: The higher the quality of external expertise for a data warehouse implementation, the greater the data quality in data warehouse.

2.1.5 Teamwork

While implementing a Data warehouse, in addition to top management commitment, three stakeholders must work together: the external consultants, the Information System (IS) staff and the end users (decision makers).

The focus on end users is very important for the success of implementing data warehouse of high data quality, because they define the business requirements (KPI, reports and dashboards) and data quality requirements as suggested by many data quality assessment methodologies. They should be also involved by data warehousing team while all the stages of the project including recipe and training.

The internal IS team plays an important role in coordinating between the different stakeholders. She is also responsible of selecting data in alignment with business requirements and giving the necessary information to external consultants.

So, every role in the project is important. But the communication between all the members is the most important. It is teamwork that allows external experts to implement a system of high data quality, fit for use for the decision makers, with the help of IS staff.

Hongjiang Xu proved that the technical factor 'Teamwork' influences Information quality in the Data warehousing success model (Xu, 2008). And reviewing quality management literature, 'Teamwork' is identified as one of the key success factors of Quality management (Cheng & Choy, 2007) .So the teamwork is important to produce an information product of high quality.

We propose then:

H5: The greater is the quality of team working on data warehouse implementation, the greater is the data quality in data warehouse.

2.1.6 Schedule

We mean by 'schedule' the planning and the time allowed to data warehouse implementation.

If it is a tight schedule in time, it pushes data warehousing team to finish quickly, and so not to give sufficient attention to data quality and not to allow sufficient time to data staging.

On another side, the implementation planning should be respected in order not to heng the project hang indefinitely in time.

In this context, Baker (Baker & Baker, 1999) and Sigal (Sigal, 1998) consider that proper planning and execution of the implementation schedule is critical to data warehouse implementation success.

So, we propose:

H6: The proper schedule for data warehouse implementation, the greater data quality in data warehouse.

2.1.7 Technology Factor

By the technology factor we mean ETL tools, Data quality tools, type of load strategy and infrastructure performance.

In data warehousing project, the ETL tools are dedicated to extract, transform and load data from data source to data warehouse. In the transformation stage, cleaning and data improvement can be done depending on transformation features and data quality features offered by the ETL tool used.

Data quality tools are also very important to get a high data quality in data warehouse. When a data quality tool is used in a data warehousing project, integrated with ETL tool, and depending on the different functionalities it offers, it allows data quality improvement.

Loading strategy has also an influence on data quality. It refers to loading type (Bulk, batch load or simple load) and

loading frequency. It impacts especially on the freshness or the timeliness dimension of data quality.
The performance of the platform behind the data warehousing process impacts the quality of data in data warehouse. It is the ability of platform used to execute the compiled code in an optimized and speed way.
H7: The technology factor supporting the data warehouse has an impact on data warehouse data quality.

2.2. Measurement model

In this section, we present the scale items to measure each latent variable in the presented research model. The items used were taken from previously validated sources and adapted to the context. (Zellal & Zaouia, A measurement model for factors influencing data quality in data warehouse, 2016)

2.2.1 Data quality measurement

As presented in literature, data quality is a multidimensional concept (Eckerson, 2006). Which means that evaluating the quality of a dataset amounts to evaluating its completeness, correctness, accuracy, consistency and so on. In fact, there are so many dimensions of data quality, and there is no general agreement on them (Fischer & Kingma, 2001). So we choose four of them which are the most frequently mentioned in literature and which constitute the focus of the majority of the authors (Emily, 1997). They are also defined as the basic set of data quality dimensions by Batini et al. (Shin, 2003) after analyzing the most important classifications of data quality dimensions. These data quality dimensions are: Timeliness, Accuracy, Consistency and Completeness.
So we'll consider these dimensions as items of measurement of data quality, but we'll give them different definitions depending on data if it is at the source or at the data warehouse.
Source Data Quality Measurement items:

- **Timeliness** indicates if data is updated (fresh) according to changes known by time in the real world.
- **Accuracy** is the measure that indicates how well and how correctly is data represented in the data base, comparing its value to the real world or to a reference data.
- **Completeness** can be defined as the measure that indicates if all the useful fields are filled.
- **Consistency** is the measure that indicates that data don't violate integrity constraints and don't conflict each other and can be considered logic referring to the business rules.

Data Warehouse Data Quality Measurement items:

- **Timeliness** indicates if data is sufficiently updated for the decision maker and for the decision task.
- **Accuracy** is the measure that indicates the correctness and precision required to make a specific decision concerned by this information
- **Completeness** can be defined as the measure that indicates complete if the users (decision makers) can deduce any necessary information they need for their decision tasks
- **Consistency** is the measure that indicates that data is not conflicting each other and not conflicting business rules and users requirements in what concerns format and content.

2.2.2 Data quality management practices

The items used to measure the data quality management adoption are:

- Definition of data quality expectations for the Decision Support System
- Definition and use of data quality dimensions accordingly
- Institution of data governance
- Agreeing to data quality standards
- Monitoring data quality performance

2.2.3 Top Management Commitment

The items we propose to measure top management commitment in the context of data quality in data warehousing project are as follows:

- Participation and support of Top management team in the data warehousing project
- Allocation of the necessary human resources to the DW project
- Allocation of the necessary financial resources to the DW project
- Attitude to change Allocation of the necessary human resources
- End user satisfaction with the changes top management decides on data quality issues
- Quality Priority: quality is treated as more important than cost and time by top management in DW project

2.2.4 External Expertise Quality

We propose the following items to measure External expertise:

1. Vendor / Consultant adequate technical support,

2. Vendor / Consultant credibleness and trustworthiness,
3. Vendor / Consultant relationship and communication with organization
4. Vendor / Consultant experience in DW projects
5. Vendor / Consultant quality training and services

2.2.5 Teamwork

We propose these items to assess Teamwork factor in our research context:

- Clear vision and elevating goals for all team members
- Competency of the team members
- Collaborative climate (sharing ideas and expertise) between the team members
- Support and recognition between the team members
- Team leadership
- Unified commitment of the team on one engagement.

2.2.6 Schedule

We propose the following items to measure the schedule factor:

- Practical Implementation Schedule
- Stable scoping of project
- Change of the planning accordingly to any change in the project scope

2.2.7 Technology Factor

To measure the technology factor we consider:

- The transformation features of ETL tool
- The data quality features of ETL and Data quality tools
- Loading Strategy
- Platform performance

3. Research Methodology

3.1 Data Collection

In order to test the hypotheses of our model, we conducted a survey in Morocco. The questionnaire targeted especially the BI specialists and end users of data warehouse in Moroccan medium and large accounts.
The questionnaire was sent via LinkedIn and by mails only to our professional network in order to avoid any fake response.
To ensure data validity and reliability, five knowledgeable individuals (i.e., 1 BI professor, 3 BI consultants and 1 BI managerial level user) completed the questionnaire before our mailing it, and their comments helped improve its quality.

3.2 Instrument Development

The measures used were taken from the established measurement model and anchored on a 5-point Likert scale, ranging from *strongly disagree* (1) to *strongly agree* (5), on which participants were asked to indicate an appropriate choice.

Data quality in data warehouse (DQDW) was assessed with the following statements:
"In your Data Warehouse, the data is updated frequently enough to allow you to make the decisions you need at the right time", "In your Data Warehouse, the data is fairly accurate and accurate for you to make your decisions", "In your Data Warehouse, you will find all the information you need to make your decisions" and "The data in your Data Warehouse is consistent, it does not represent a conflict between them, or a conflict with business rules"
Data quality in source systems (DQSS) was assessed by the following:
"In your Operational Information System (OIS), the data are updated according to their variation in reality", "In your OIS, the data is precise and represent exactly their respective elements in reality" You need it for your daily operational work you find it in your OIS"," The data in your OIS are consistent and do not represent a contradiction between them "

The adoption of data quality management practices (DQMP) was assessed with these statements:
"Expectations of end users in terms of data quality in the decision-making system are well identified and documented", "Data quality measures are well defined and documented to measure the achievement of end-user expectations in terms of Data quality "," Your company establishes good data governance (Definition of processes, roles and responsibilities for data quality) "," Standards and good practices of data quality are your reference in any step relating to data: data collection, transformation, updating ... ", "You set up a continuous improvement system to measure the achievement of the objectives in terms of data quality "
The Top Management commitment (TM) was assessed using the following statements:

"Top Management supports the BI project by actively participating in its management", "The TM is ready to allocate all the human resources needed for the project BI", "The TM is ready to allocate all the financial resources it needs to successfully implement a DW with a good quality of data”, “TM is willing to change existing work and procedures to improve data quality”, “End users are satisfied with the changes that Top Management decides on issues of data quality in the data warehousing project”, “Quality is considered by the TM as more important than the cost and time in the data warehousing project”

The External Expertise (EE) was assessed by these statements:

"Your service provider has the right technical support", "Your service provider is credible and trustworthy", "Your service provider has a good relationship with your organization (Project team and decision makers)", "Your service provider has experience in the field Decision-making", "Your provider offers high quality training and services"

The Teamwork (Team) was assessed using the following statements:

"The whole team has a clear vision of the decision-making project (its objectives, deadlines, sources, users ...)", "The members of the project team are competent", "There is a collaborative climate between members Project team (sharing of ideas, experience and knowledge)", "There is help and recognition among team members project"," there is good leadership (direction) Team", "The project team has a unified commitment on which all members agree"

The Schedule (Sch) factor was assessed by these 3 statements:

"The implementation planning of the decision-making system in your organization is practical and reasonable", "The decision-making perimeter is stable throughout the project" and "the implementation planning is reviewed and modified every time perimeter of the project is modified"

The technology factor (TF) was assessed by the following statements:

"You are satisfied with the data transformations offered by the ETL used to build your Data Warehouse", "You are satisfied with the quality improvements of your tools (ETL or QD tool), For example: real-time cleaning, verification of data according to business rules", "You are satisfied with the data loading strategy (loading flow schedules), ie it does not impact Data quality in your Data Warehouse" and "You are satisfied with the performance of the infrastructure that supports your Data Warehouse (High Availability, Speed of Code Processing)"

The construct reliabilities of the measures as assessed by Cronbach alphas are high (greater than 0.7) compared with recommended values in the literature (Nunnally, 1978), as shown in the following figure.

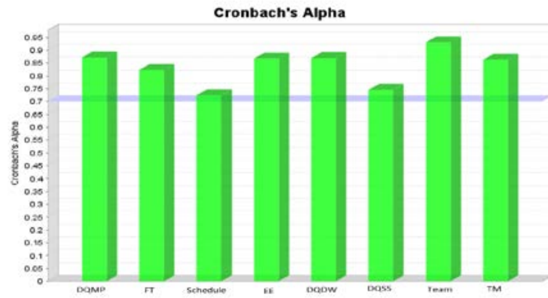

Figure 2: Measures Cronbach's Alpha

3.3 The sample

The overall response rate was 33%. In total, we received 80 individual responses. The responses were received from diverse industries: Banks and Assurances (25%), Telecommunications (12,5%), Industry (11,3%), Transport (8,8%), Finance (6,3%), Consulting (6,3%) and Health sector (3,7%) …

The respondents’ positions in the organizations vary from junior employee to top manager. Most of them are senior employees (40%). The majority of participants work in Information Systems direction (75%) and the other minority is spread over different directions such as Business operations, marketing and control management. More than 90% of the participants have an IT background, thing that allowed them to understand and rate the questionnaire statements easily.

The sample includes small (19%), medium-sized (36%) and large firms (45%). 47% of them have implemented their data warehouse more than 5 years ago.

3.4 Data Analysis

A structural equation modeling (SEM) technique was used to examine the relationships among the constructs.

SEM is a powerful technique, widely used in the behavioral sciences that can combine complex path models with latent variables (Hox & Bechger).

There are two main approaches: PLS (Partial Least Squares) and covariance-based SEM. The PLS approach was chosen for its capability to accommodate small-sized samples (Chin, 1998).

Additionally, PLS recognizes two components of a casual model: the measurement model and the structural model.

The measurement model consists of relationships between the latent variables and the measures underlying each construct. PLS method allows the demonstration of the construct validity of the research instrument (i.e. how well the instrument measures what it purports to measure). The two main dimensions are the convergent validity and the discriminant validity. The convergent validity (also known as the composite reliability) assesses the extent to which items on a scale are theoretically related. It reflects if the

measures of constructs that theoretically should be related to each other are, in fact, observed to be related to each other. And the discriminant validity shows if the measures of constructs that theoretically should not be related to each other are, in fact, observed to not be related to each other.
On the other hand, the structural model provides information on how well the hypothesized relationships predict the theoretical model.

PLS software e.g. Smart PLS 3.0 (the software we used for our PLS analysis), provides the squared multiple correlations (R2) for each endogenous construct in the model and the path coefficients. The coefficient of determination R2 indicates the proportion of variance (%) in the dependent variable that can be explained by the independent variable in the model while the path coefficients (β) indicate the strengths of relationships between constructs (Chin, 1998). Chin (1998) notes that both the β and the R2 are sufficient for analysis, and β values between 0.20 and 0.30 are adequate for meaningful interpretations.

3.4.1Assessment of the measurement model

Smart PLS computed the composite reliability of each construct and also showed the item loading (Figure below).

	Outer loadings	Cronbach's Alpha	Construct reliability
DQMP_1	0.794	0.872	0.662
DQMP_2	0.895		
DQMP_3	0.861		
DQMP_4	0.829		
DQMP_5	0.672		
Schedule_1	0.840	0.725	0.642
Schedule_2	0.784		
Schedule_3	0.778		
EE_1	0.726	0.867	0.654
EE_2	0.859		
EE_3	0.866		
EE_4	0.826		
EE_5	0.758		
DQDW_Coherence	0.859	0.869	0.719
DQDW_MAJ	0.798		
DQDW_Precision	0.874		
DQDW_comp	0.860		
DQSS_Coherence	0.810	0.746	0.565
DQSS_Comp	0.787		
DQSS_MAJ	0.664		
DQSS_Precision	0.738		
TEAM_1	0.780	0.931	0.744
TEAM_2	0.833		
TEAM_3	0.907		
TEAM_4	0.912		
TEAM_5	0.826		
TEAM_6	0.908		
FT_1	0.848	0.823	0.658
FT_2	0.858		
FT_3	0.847		
FT_4	0.677		
TM_1	0.762	0.863	0.589
TM_2	0.824		
TM_3	0.812		
TM_4	0.644		
TM_5	0.797		
TM_6	0.754		

Figure 3: Questionnaire's measures, Loadings, Cronbach alphas, and Construct reliabilities

Fornell and Larcker (1981) note that item loadings and composite reliabilities greater than 0.7 are considered adequate (Fornell & Larcker, 1981).
In assessing the discriminant validity, the square root of the average variance extracted (AVE) for each construct, which provides a measure of the variance shared between a construct and its indicators, is checked. Fornell and Larcker (1981) and Chin (1998) recommend AVE values of at least 0.50 and that the square root of AVE should be larger than off-diagonal elements. The results in the figure below indicate that in no case was there any correlation between the constructs greater than the squared root of AVE (the leading diagonal). This suggests that the study's measures are distinct and unidimensional. Which means that the convergent and discriminant validity of the data are psychometrically adequate for this study.

	DQMP	FT	Schedule	EE	DQDW	DQSS	Team	TM
DQMP	**0.814**							
FT	0.649	**0.811**						
Schedule	0.506	0.579	**0.801**					
EE	0.446	0.651	0.454	**0.809**				
DQDW	0.586	0.753	0.533	0.594	**0.848**			
DQSS	0.279	0.45	0.42	0.429	0.62	**0.752**		
Team	0.47	0.594	0.523	0.433	0.68	0.438	**0.863**	
TM	0.589	0.394	0.493	0.266	0.491	0.4	0.439	**0.768**

Figure 4: Fornell-Larcker Criterion

3.4.2Assessment of the structural model

The paths coefficients (β) and the R2 were generated by Smart PLS 3.0. Values are shown in the following figure. The R2 is 0.73, which suggests that the contingency factors explained 73% of the variance in the *Data Warehouse Data Quality* construct. This value is considered strong effect size (Moore, Notz, & Flinger, 2013).

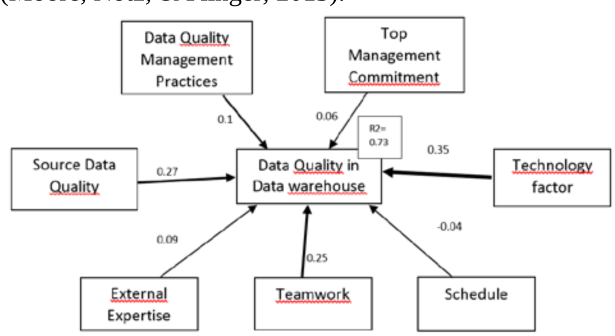

Figure 5: The Smart PLS Graph results for the research model

4. Discussion and Conclusion

The objective of this study was to find the most critical factors for data quality in data warehouse while data warehouse implementation.

This empirical study reveals that the contingent factors that we included in our research model were able to explain 73% of the variance of data quality in data warehouse.
The most critical factor, according to our survey conducted in Morocco, is the technology factor (β= 0.35). That means that a special attention should be given to the choice of ETL and data quality tools while the data warehouse implementation. The platform performance and loading strategy are also very important.
The second most critical factor is data quality in source systems (β= 0.27). This means that if an organization needs a data warehouse of good data quality, it should start by improving the data quality in the source systems.
Another critical factor that has been supported by the survey is the teamwork (β= 0.25). This can be considered as a key success factor for data quality in data warehouse. It requires a good leadership, competent team members and a sharing and recognition spirit.
The factor “Adoption of data quality management practices” has been only moderately supported by the collected data (β= 0.102), while the other factors “Schedule”, “Top Management commitment” and “External Expertise” was not really supported by collected data.

References

[1] Alhyasat, E. B. (2013). Data Warehouse Success and Strategic Oriented Business Intelligence: A Theoretical Framework. Journal of Management Research, 5(3).
[2] Amit, R., & Emilie, Y. (1999). Key Issues in Achieving Data Quality and Consistency in Data Warehousing among Large Organizations in Australia. International Conference on System Sciences. Hawaii.
[3] Baker, S., & Baker, K. (1999). The best little warehouse in business. Journal of Business Strategy, 32-37.
[4] Briggs. (2002). A Critical Review of Literature on Data Warehouse Systems Success/Failure. Journal of Data Warehousing, 41, 1-20.
[5] Cheng, E., & Choy, P. (2007). Measuring Success Factors of Quality Management in the Shipping Industry. Palgrave Journals, 234-253.
[6] Chin, W. (1998). Issues and opinion on structural equation modeling. MIS Quarterly, 22(1).
[7] Dale, & Duncalf. (1985). Quality-related decision making: A study in six British companies. nternational Journal of Operations and Production Management, 5(1), 15-25.
[8] Eckerson, W. (2006). Data quality and the bottom line: Achieving business success through the commitment to high quality data. 101 Communications LLC.
[9] Emily, K. (1997, October). Dirty data challenges warehouses. Software Magazine.
[10] Fischer, C., & Kingma, B. (2001). Criticality of data quality as exemplified in two disasters. Information & Management, 109-116.
[11] Fornell, C., & Larcker, D. (1981). Evaluating structural equations models with unobservable variables and measurement error. Journal of Marketing Research, 8(1), 39-50.
[12] Haley, Watson, H., & Barbara. (1997). Data warehousing: A framework and survey of current practices. Journal of Data Warehousing, 10-17.
[13] Hox, J., & Bechger, T. (s.d.). An Introduction to Structural Equation Modeling. Family Science review, 354-373.
[14] Ifinedo, P. (2008). Impacts of business vision, top management support, and external expertise on erp success. Business Process Management Journal, 14(4), 551-568.
[15] Inmon, W. H. (1992). Building the Data Warehouse. John Wiley & Sons,.
[16] Lee, Y.-S., Lee, D.-M., & and Suh, C.-K. (2001). Factors Affecting Implementation of Data Warehouse. PACIS, (p. 47).
[17] Moore, D. S., Notz, W. I., & Flinger, M. A. (2013). The basic practice of statistics. New York: Freeman and Company.
[18] Nunnally, J. (1978). Psychometric Theory. New York: McGraw-Hill.
[19] Orr, K. (1998). Data quality and systems theory. 41, pp. 66-71.
[20] Shankaranarayanan, G., & Cai, Y. (2006). Supporting data quality management in decision-making. Decision Support Systems, 302-317.
[21] Shin, B. (2003). An Exploratory Investigation of System Success Factors in Data Warehousing. Journal of the Association for Information Systems, 4(1).
[22] Sigal, M. (1998). A common sense of development strategy. Communications of the ACM , 42-43.
[23] Thong, J., Yap, C.-s., & Raman, K. (1996). Top management support, external expertise and information systems implementation in small businesses. Information Systems Research, 7(2).
[24] Wang, E., Shih, S. P., Jiang, J., & Klein, G. (1996). The consistency among facilitating factors and ERP implementation success: A holistic view of fit. Journal of Systems and Software, 7(2).
[25] WEBER, K., OTTO, B., & OSTERLE, H. (2009). One size does not fit all — a contingency approach to data governance. ACM Journal of Data and Information Quality.
[26] Wixom, B., & Watson, H. (2001). An empirical investigation of the factors affecting data warehousing success. MIS Quarterly, 25(1), 17-41.
[27] Xu, H. (2008). A Structural Model of Data Warehousing Success. Scholarship and Professional Work Business.
[28] Yeoh, W., & Koronios, A. (2010). Critical success factors for business intelligence. Journal of computer information systems, 23-32.
[29] Zellal, N., & Zaouia, A. (2015). An exploratory investigation of Factors Influencing Data Quality in Data Warehouse. WCCS. Marrackech: IEEE.
[30] Zellal, N., & Zaouia, A. (2016). A measurement model for factors influencing data quality in data warehouse. CIST. Tangier: IEEE.

Zaouia Abdellah is a professor and researcher in Science and Information Technology Management at the National Institute of Posts and Telecommunications (Rabat, Morocco). After twenty years of experience in management and IT consulting in the public sector and the private sector, he devotes his scientific research and PhD thesis supervizing to the field of transactional information and decision-making systems. He Provides also training, supervises engineering projects in the field of ERP, Business Intelligence and Big Data with case studies applied to business management and computer simulations.

Nouha Zellal is a PhD Student in Information Technology Management at the National Institute of Posts and Telecommunications and an IT consultant.
After she got her engineering degree in 2013, she has shared her time between BI, CRM and Analytics consulting and scientific research to give her researches a practical side that benefits from the literature review and field experience.