SPREADSHEET AS A MEANS TO SUPPORT NONCONFORMITY ANALYSIS – DATA INTEGRITY IN THE TOOL

SZCZĘŚNIAK Bartosz

Institute of Production Engineering,
Faculty of Organisation and Management, 41-800 Zabrze, ul. Roosevelta 26-28,
Silesian University of Technology
E-mail: bartosz.szczesniak@polsl.pl

Abstract: The solution discussed in the article is an example of a tool created in accordance with the author's signature concept of Microtools Based on the Relational Data Model to support the nonconformity analysis process regarding rolled products. The study opens with an explanation of the data structure implemented in the tool. The entities thus identified as well as their attributes and interlinks between them have been depicted using an adequate Entity Relationship Diagram. What has also been discussed in the manner in which the proposed data structure has been implemented in a spreadsheet. Further paragraphs of the publication describe the solutions applied to ensure integrity of the data stored in the tool. The conditions to be met by the data as well as individual mechanisms created to verify them have been described. All the solutions proposed in the paper have been developed using standard spreadsheet features and components, without involving any additional code created in any programming language.

Keywords: information processes, process improvement, relational data model, MiRel.

1 Introduction

Functioning of contemporary businesses is inextricably linked with the notion of information. Processes where information is generated, collected, stored, processed, transferred, rendered, interpreted or used are referred to as information processes [4, 5]. Different kinds of IT tools are commonly used to support their implementation. There is a group of computer programs used in nearly every organisation and familiar to nearly every employee, namely spreadsheets. They enable relatively quick creation of tools matching the given company's current needs, which prove particularly useful in situations when application of other IT solutions is impossible. Using spreadsheets to support the functioning of organisations, as proposed in the literature of the subject, is a concept which encompasses a very wide range of applications. The diverse solutions proposed are, among other spheres, related to finance management, controlling [2, 3, 14], sales, marketing [2, 14, 15] or quality management [2, 10, 11].

The tools thus created may be developed in an intuitive manner, without any openly predefined data model. They may also be preceded by a thorough problem analysis leading to formulation of an adequate relational data model which – once it has been implemented in the spreadsheet – provides grounds for the tool to be built [6, 12, 13]. And this is exactly the kind of procedure assumed as a basis for the author's signature concept of Microtools Based on the Relational Data Model, abbreviated to MiRel [7]. An example of a tool conforming with this concept is one which supports the process of product nonconformity analysis. Individual solutions which make it possible to develop the reports assumed as the process deliverables have been discussed in the author's previous publication [8]. Further sections of this article describe the data structure implemented in the tool in question as well as the solutions applied to ensure integrity of the data entered into and stored in the tool.

2 Data structure applied in the tool

In the development of the tool addressed in the paper, it was originally proposed that it should comprise such entities as Order, Nonconformity Card, Cause, Decision, Production Line, Control, Cause on a Card, Nonconformity for Cause on a Card and Decision for Cause on a Card. The structure of the above entities along with the attributes they comprise has been depicted in a diagram conforming with the CASE Method [1], as provided in Figure 1.

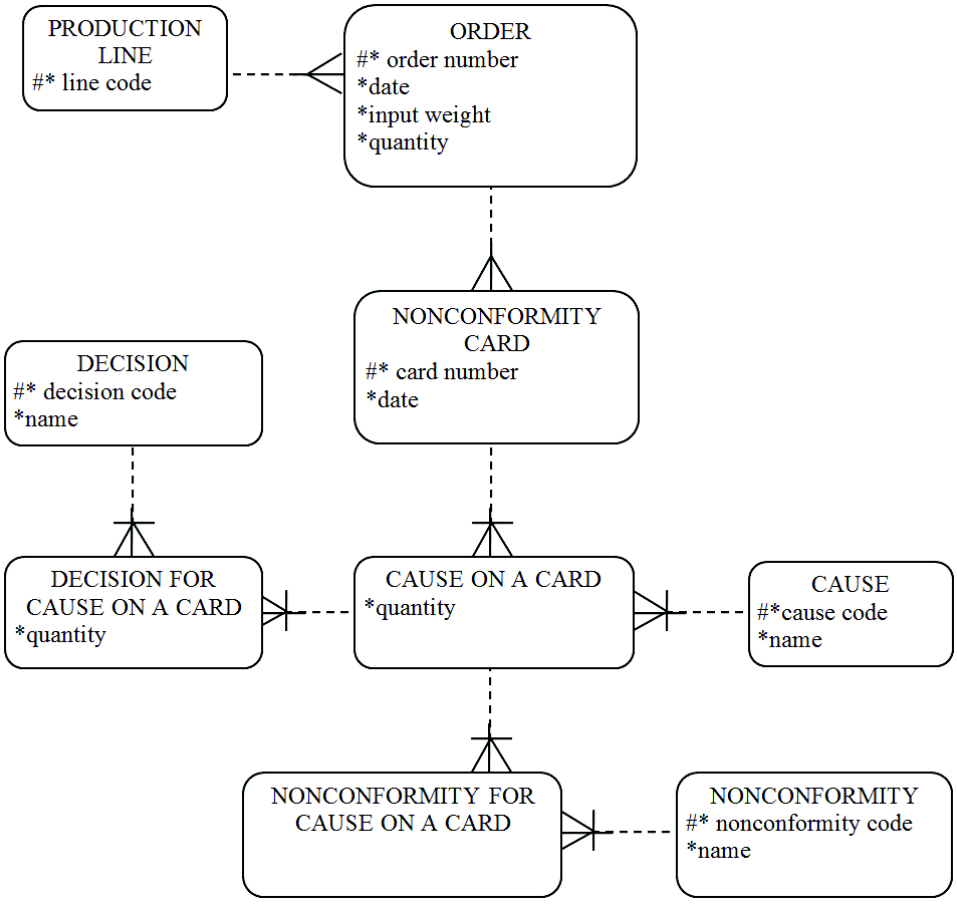
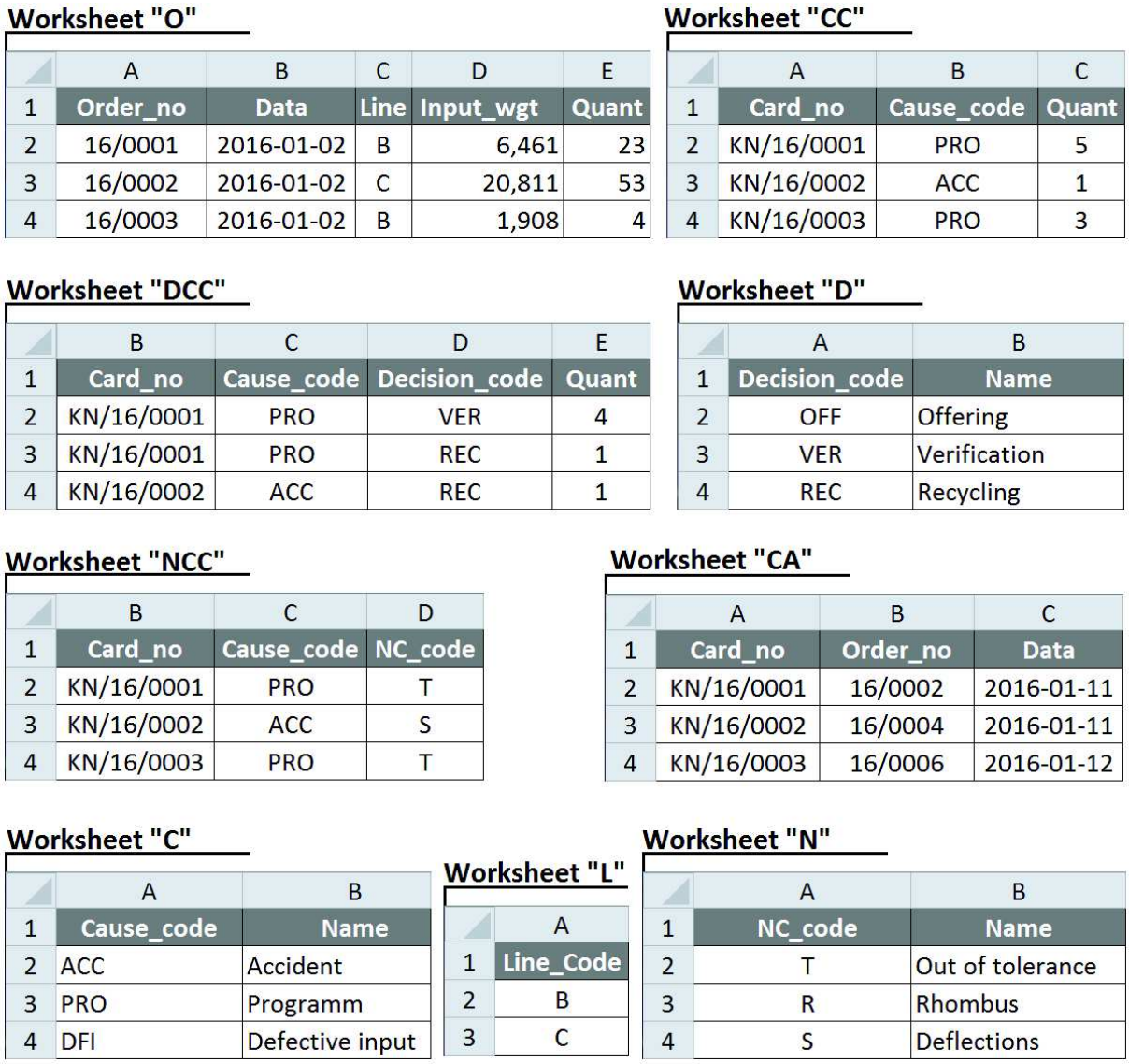


Fig. 1. Structure of entities in the tool discussed

The MiRel concept is based on an assumption that entities can be represented in a spreadsheet in two ways [9]. In this case, all the identified entities have been implemented as traditional tables, as it is handled in classical relational databases. Each table has been entered into a separate worksheet. The worksheet names are, at the same time, table names. Their arrangement has been shown in Figure 2.



Worksheet "O"

	A	B	C	D	E
1	Order_no	Data	Line	Input_wgt	Quant
2	16/0001	2016-01-02	B	6,461	23
3	16/0002	2016-01-02	C	20,811	53
4	16/0003	2016-01-02	B	1,908	4

Worksheet "CC"

	A	B	C
1	Card_no	Cause_code	Quant
2	KN/16/0001	PRO	5
3	KN/16/0002	ACC	1
4	KN/16/0003	PRO	3

Worksheet "DCC"

	B	C	D	E
1	Card_no	Cause_code	Decision_code	Quant
2	KN/16/0001	PRO	VER	4
3	KN/16/0001	PRO	REC	1
4	KN/16/0002	ACC	REC	1

Worksheet "D"

	A	B
1	Decision_code	Name
2	OFF	Offering
3	VER	Verification
4	REC	Recycling

Worksheet "NCC"

	B	C	D
1	Card_no	Cause_code	NC_code
2	KN/16/0001	PRO	T
3	KN/16/0002	ACC	S
4	KN/16/0003	PRO	T

Worksheet "CA"

	A	B	C
1	Card_no	Order_no	Data
2	KN/16/0001	16/0002	2016-01-11
3	KN/16/0002	16/0004	2016-01-11
4	KN/16/0003	16/0006	2016-01-12

Worksheet "C"

	A	B
1	Cause_code	Name
2	ACC	Accident
3	PRO	Programm
4	DFI	Defective input

Worksheet "L"

	A
1	Line_Code
2	B
3	C

Worksheet "N"

	A	B
1	NC_code	Name
2	T	Out of tolerance
3	R	Rhombus
4	S	Deflections

Fig. 2. Representation of entities in a spreadsheet

3 Solutions ensuring data integrity

An important problem tackled when developing a tool in a spreadsheet is how to create solutions that enable validation of the data being entered. In this respect, the correctness may pertain to different aspects, one of which is the conformity of foreign key values with the values of simple primary keys contained in source tables. In order to ensure this kind of conformity in columns, where foreign key values are entered, a *data validation* mechanism has been implemented, so that one can enter values selected from a drop-down list. The data sources for these lists are specific named worksheet ranges, as defined in the name manager. These ranges have been

dynamically defined by means of suitable functions. The named ranges have been listed along with their definitions and target columns in Table 1.

Table 1. Named ranges used as data sources for data validation

Name	Reference	Target table	Target column
Card_Number	=OFFSET(CA!$A$2;0;0; COUNTA(CA!$A$2:$A$1001);1)	CC	Card_no
Cause_Code	=OFFSET('C'!$A$2;0;0; COUNTA('C'!$A$2:$A$11);1)	CC	Cause_code
Decision_Code	=OFFSET(D!$A$2;0;0; COUNTA(D!$A$2:$A$11);1)	DCC	Decision_code
Line_Code	=L!$A$2:$A$3	O	Line
Nonconformity_Code	=OFFSET(N!$A$2;0;0; COUNTA(N!$A$2:$A$11);1)	NCC	NC_code
Order_Number	=OFFSET(O!$A$2;0;0; COUNTA(O!$A$2:$A$1001);1)	CA	Order_no

The other data integrity problem one must solve is connected with conformity between foreign key values and values of corresponding composite primary keys. In the tool in question, this is the case of the composite primary key in table "CC". This key comprises attributes "Card_no" and "Cause_code", and it migrates to tables "DCC" and "NCC". The combinations of attributes "Card_no" and "Cause_code" entered into these tables must already exist in table "CC". A solution proposed in order to ensure this conformity is to add auxiliary columns in worksheets which contain the three aforementioned tables.

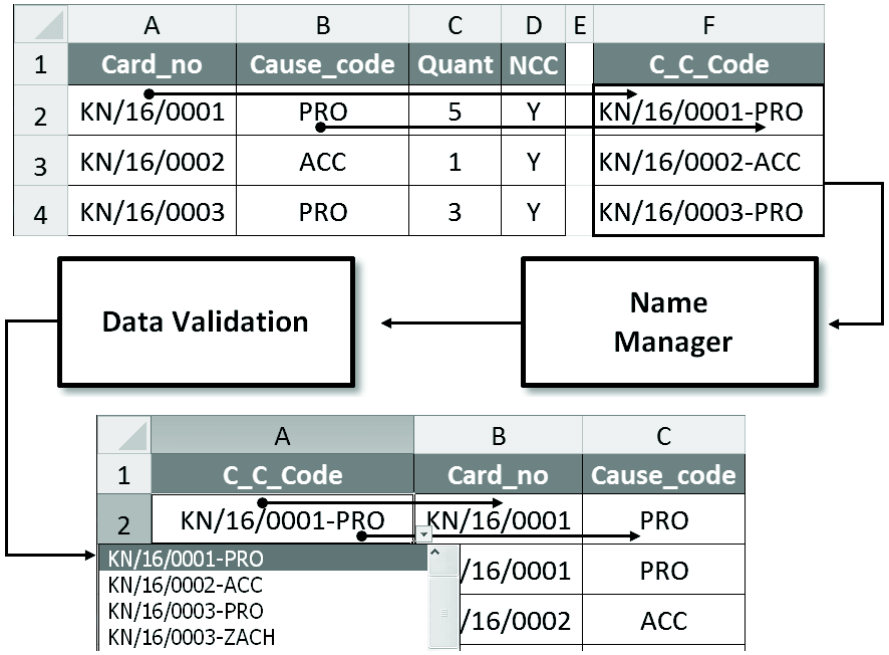


Fig. 3. Schematic representation of the solution ensuring conformity of key values in tables "CC", "NCC" and "DCC"

In worksheet “CC”, in the added column, a code is created as a combination of attributes “Card_no” and “Cause_code” using an appropriate formula. The named worksheet range that contains this code becomes the data source for the drop-down list used by the data validation mechanism. In worksheet “DCC” or “NCC”, in the added column, the user enters a value by selecting it from a list of permissible existing combinations of attributes “Card_no” and “Cause_code”. The value thus entered is broken down into values of individual attributes by means of suitable formulas. This solution has been schematically illustrated in Figure 3. The formulas applied in it have been depicted in Figure 4 using the example of tables “CC” and “DCC”.

Worksheet "CC"

	A	B	C	D	E	F
1	Card_no	Cause_code	Quant	NCC	FA	C_C_Code
2	KN/16/0001	PRO	5	Y		KN/16/0001-PRO
3	KN/16/0002	ACC	1	Y		KN/16/0002-ACC

Worksheet "DCC"

	A	B	C	D	E
1	C_C_Code FB	Card_no FC	Cause_code	Decision_code	Quant
2	KN/16/0001-PRO	KN/16/0001	PRO	VER	4
3	KN/16/0001-PRO	KN/16/0001	PRO	REC	1

FA	IF(A2<>"";CONCATENATE(A2;"-";B2);"")
FB	IF(A2<>"";LEFT(A2;SEARCH("-";A2)-1);"")
FC	IF(A2<>"";MID(A2;SEARCH("-";A2)+1;LEN(A2)-SEARCH("-";A2));"")

Fig. 4. Formulas applied in the solution to ensure conformity of keys in tables “CC” and “DCC”

Another issue that one must probably address in terms of data integrity is the unrepeatability of primary key values. In order to solve it, auxiliary tables have been added to individual tables, where – in each row – the algorithm counts the number of instances of the primary key from the given row in the table. In the event that a value repeats itself, the cells which contain it become highlighted by conditional formatting. An example of this solution with reference to table “CC” has been illustrated in Figure 5.

Conditional Formatting Rule		
Rule	Format	Applies to
Formula: =$H2>1	AaBbCcYyZz	=$A$2:$B$3001

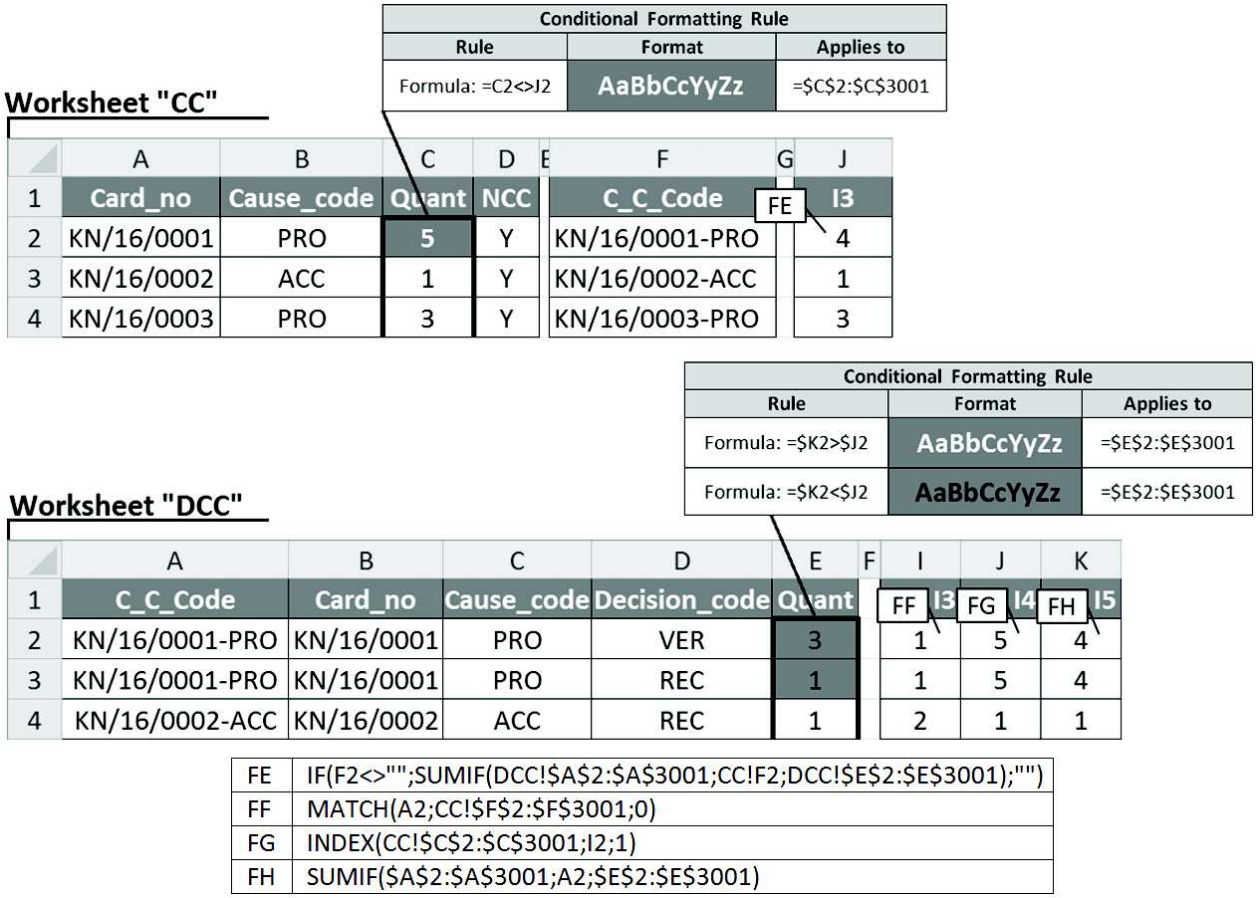
	A	B	C	D	G	H
1	Card_no	Cause_code	Quant	NCC	FD	I1
2	KN/16/0001	PRO	5	Y		1
3	KN/16/0002	ACC	1	Y		1
4	KN/16/0003	PRO	3	Y		2
5	KN/16/0003	PRO	3	Y		2

FD	COUNTIFS($A$2:$A$3001;A2;$B$2:$B$3001;B2)

Fig. 5. Solution applied to verify primary key uniqueness in table “CC“

The solutions discussed above have all concerned uniqueness of primary keys and conformity between their values and those of foreign keys. Another condition pertains to the relationship

between the data entered into tables “CC” and “DCC”. The number of pieces assigned to a specific cause on the given card in table “CC” must correspond to the sum of pieces assigned to different decisions for the given cause and card in table “DCC”. In order to verify this condition, auxiliary columns have been added both in worksheet “CC” and in worksheet “DCC”. Based on the values to be found in the auxiliary columns, in both tables, cells in columns “Quant” whose values do not match one another are highlighted using conditional formatting. The added auxiliary columns along with the formulas applied and the conditional formatting rules have all been depicted in Figure 6.



	A	B	C	D
1	Card_no	Cause_code	Quant	NCC
2	KN/16/0001	PRO	5	Y
3	KN/16/0002	ACC	1	Y
4	KN/16/0003	PRO	3	Y

	F	J
1	C_C_Code	I3
2	KN/16/0001-PRO	4
3	KN/16/0002-ACC	1
4	KN/16/0003-PRO	3

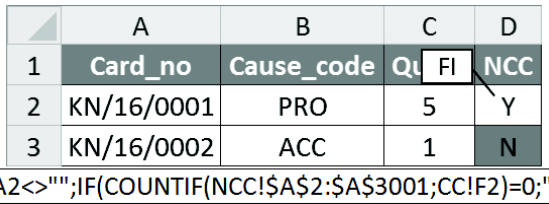
	A	B	C	D	E
1	C_C_Code	Card_no	Cause_code	Decision_code	Quant
2	KN/16/0001-PRO	KN/16/0001	PRO	VER	3
3	KN/16/0001-PRO	KN/16/0001	PRO	REC	1
4	KN/16/0002-ACC	KN/16/0002	ACC	REC	1

	I	J	K
1	I3	I4	I5
2	1	5	4
3	1	5	4
4	2	1	1

FE	IF(F2<>"";SUMIF(DCC!$A$2:$A$3001;CC!F2;DCC!$E$2:$E$3001);"")
FF	MATCH(A2;CC!$F$2:$F$3001;0)
FG	INDEX(CC!$C$2:$C$3001;I2;1)
FH	SUMIF($A$2:$A$3001;A2;$E$2:$E$3001)

Fig. 6. Arrangement of formulas and formatting rules of the data validating mechanism for tables “CC” and “DCC”

Another verification mechanism applied in the tool is intended to check whether each combination of values of attributes “Card_no” and “Cause_code” that appears in table “CC” also appears at least once in table “NCC”. In order to test this condition, the “NCC” auxiliary columns has been added to table “CC”. Both the column and the formula applied have been shown in Figure 7.



	A	B	C	D
1	Card_no	Cause_code	Qu	NCC
2	KN/16/0001	PRO	5	Y
3	KN/16/0002	ACC	1	N

FI	=IF(A2<>"";IF(COUNTIF(NCC!$A$2:$A$3001;CC!F2)=0;"N";"Y");"")

Fig. 7. Additional column and formula of the mechanism verifying conformity of data in tables “CC” and “NCC”

The cells of auxiliary column “NCC” which contain the “N” value have been highlighted by the conditional formatting mechanism. This implies that an appropriate combination of “Card_no” and “Cause_code” has not been entered into table “NCC” yet.

The last of the solutions proposed verifies whether values of all attributes have been entered in all the tables previously introduced. What has been used in this solution is the conditional formatting mechanism based on the added auxiliary column. An example of such a solution assumed for table “CC” has been illustrated in Figure 8.

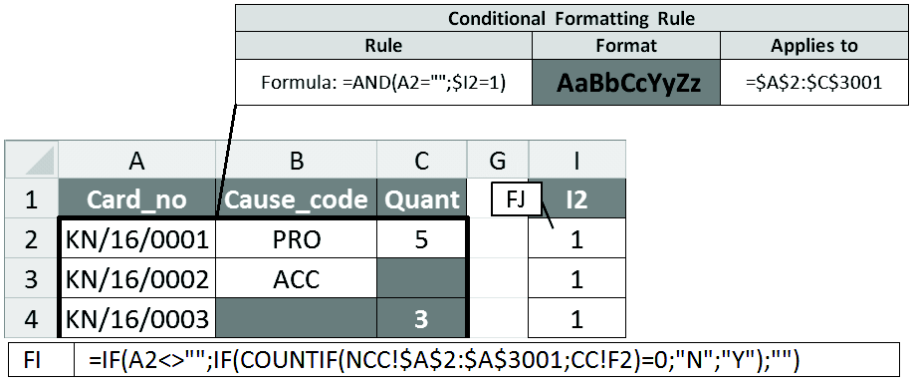


Fig. 8. Model solution verifying whether all attributes have been entered

Conclusion

The solutions proposed in the paper have illustrated that, in the tool developed in accordance with the concept of Microtools Based on the Relational Data Model to support the nonconformity analysis process, data integrity may be successfully achieved by means of built-in spreadsheet components and functions. For purposes of the tool subject to analysis, the following conditions have been identified as those which must be met while data are entered:

- values of foreign keys must conform with values of simple primary keys contained in source tables,
- values of foreign keys must conform with values of composite primary keys contained in source tables,
- values of primary keys must be unique,
- value of the “Quant” attribute assigned to a specific cause and card in table “CC” must conform with the sum of values of the “Quant” attribute assigned to different decisions for the given cause and card in table “DCC”,
- each combination of values of attributes “Card_no” and “Cause_code” contained in table “CC” must appear at least once in table “NCC”,
- values of all other attributes must be present in records where a value of any of the attributes has been entered.

The components of the MS Excel spreadsheet which have been used to create solutions that verify if the above conditions are met include data validation, name manager and conditional formatting. Some standard built-in function have also been applied, including: OFFSET(), COUNTA(), COUNTIF IF(); CONCATENATE(), LEFT(), SEARCH(), MID(), LEN(), COUNTIFS(), SUMIF(), MATCH() and INDEX(). The data validation solutions discussed above make it significantly easier to use the tool analysed in the paper.

PŘEKRAČUJEME HRANICE
PRZEKRACZAMY GRANICE
2014—2020

EVROPSKÁ UNIE / UNIA EUROPEJSKA
EVROPSKÝ FOND PRO REGIONÁLNÍ ROZVOJ
EUROPEJSKI FUNDUSZ ROZWOJU REGIONALNEGO

Projekt jest współfinansowany ze środków Europejskiego Funduszu Rozwoju Regionalnego
oraz z budżetu państwa RP „Przekraczamy Granice"

References

[1] Berker R., CASE Method, Modelowanie związków encji, Wydawnictwa Naukowo Techniczne. Warszawa 1996

[2] Carlberg C., Microsoft Excel 2007 PL. Analizy biznesowe. Rozwiązania w biznesie. Wydanie III. Helion. Gliwice 2009

[3] Knight G.: Excel. Analiza danych biznesowych. Helion. Gliwice 2006

[4] Oleński J., Ekonomika informacji. Podstawy, PWE , Warszawa 2001

[5] Oleński J., Ekonomika informacji. Metody, PWE, Warszawa 2002

[6] Szczęśniak B., Wolniak R., Improving the process of car operating costs accounting using a spreadsheet-based tool – a case study, 16th International Multidisciplinary Scientific GeoConference SGEM 2016, www.sgem.org, SGEM2016 Conference Proceedings, ISBN 978-619-7105-58-2 / ISSN 1314-2704, June 28 - July 6, 2016, Albena, Bulgaria, Book2 Vol. 1, 239-246 pp, DOI: 10.5593/SGEM2016/B21/S07.031

[7] Szczęśniak B., Mikronarzędzia bazujące na relacyjnym modelu danych a rozwój specjalizacji inteligentnych, Systemy Wspomagania w Inżynierii Produkcji -Metody i Narzędzia Inżynierii Produkcji dla Rozwoju Inteligentnych Specjalizacji, 4(16)/2016 pp.121-131

[8] Szczęśniak B., Koncepcja Mikronarzędzi Bazujących na Relacyjnym Modelu Danych we wspomaganiu procesu analizy niezgodności wyrobów walcowanych, Systemy Wspomagania w Inżynierii Produkcji - Sposoby i środki doskonalenia produktów i usług na wybranych przykładach, volume 6, issue 8, 2017 pp. 59 – 72

[9] Szczęśniak B., Microtools Based on the Relational Data Model – representation of entities in a spreadsheet, 17th International Multidisciplinary Scientific GeoConference SGEM 2017, www.sgem.org, SGEM2017 Conference Proceedings, Informatics, Geoinformatics and Remote Sensing, ISBN 978-619-7408-01-0 / ISSN 1314-2704, June 29 - July 5, 2017, Albena, Bulgaria, Vol. 17, Issue 21, pp. 447-454

[10] Szczęśniak B., Zastosowanie arkusza kalkulacyjnego do wspomagania metody ABC, Zeszyty Naukowe Politechniki Śląskiej, seria Organizacja i Zarządzanie z.50, 2010, pp. 23-33

[11] Szczęśniak B., Molenda M., Spreadsheet application supporting the x-r control chart, Conference Proceedings - 22th Conference Modern Mathematical Methods in Engineering (3mi) , June 3 - 5, 2013 Horní Lomná, Czech Republic, pp. 128-134

[12] Szczęśniak B., Koncepcja zastosowania arkusza kalkulacyjnego do wspomagania tworzenia dokumentów w procesie produkcji taśm blachy. w: Systemy wspomagania w inżynierii produkcji. Innowacyjność, jakość, zarządzanie. Monografia. Red. Witold Biały, Katarzyna Midor. Gliwice : Wydawnictwo. PA NOVA, 2013, pp. 133-151

[13] Szczęśniak B., Arkusz kalkulacyjny w doskonaleniu procesu układania planu zajęć w szkole specjalnej, w: Komputerowo zintegrowane zarządzanie, Tom II. Pr. zb. pod. red. Ryszarda Knosali. Oficyna Wydawnicza Polskiego Towarzystwa Zarządzania Produkcją, Opole 2010, pp. 525-537

[14] Wilczewski S., Wrzód M., Excel 2007 w firmie. Controlling, finanse i nie tylko. Helion. Gliwice 2008

[15] Winston Wayne L., Microsoft Excel 2010. Data Analysis and Business Modeling. Third Edition. Microsoft Press. Washington 2011

ARKUSZ KALKULACYJNY WE WSPOMAGANIU ANALZY NIEZGODOŚCI – INTEGRALNOŚĆ DANYCH W NARZĘDZIU

Abstract: Przedstawione w artykule narzędzie stanowi przykład narzędzia stworzonego zgodnie z autorską koncepcją Mikronarzędzi Bazujących na Relacyjnym Modelu Danych do wspomagania procesu analizy niezgodności wyrobów walcowanych. W opracowaniu w pierwszej kolejności omówiono zastosowaną w narzędziu strukturę danych. Zidentyfikowane encje, ich atrybuty, oraz występujące pomiędzy nimi powiązania przedstawiono za pomocą odpowiedniego diagramu związków encji. Przedstawiono także sposób implementacji struktury danych w narzędziu w arkuszu kalkulacyjnym. W dalszej części publikacji opisano rozwiązania zastosowane w celu zapewnienia integralności przechowywanych w narzędziu danych. Określono warunki, jakie dane te muszą spełniać oraz opisano mechanizmy stworzone w celu ich weryfikacji. Wszystkie zaproponowane rozwiązania zostały stworzone przy wykorzystaniu wbudowanych funkcji i elementów arkusza kalkulacyjnego bez konieczności tworzenia kodu w jakimkolwiek języku programowania.

Słowa kluczowe: procesy informacyjne, doskonalenie procesów, relacyjny model danych, MiRel