Multidimensional pairwise comparison – about human-oriented science regarding artificial intelligence and value surveys

Anikó Balogh, László Pitlik, Ferenc Szani, Apertus Nonprofit Ltd.

Abstract: Human thinking is intuitive but from the point of view of logic it is mostly inconsistent. Therefore value surveys always present the analytical problem to explore the quality and quantity of inconsistences behind the averages of opinions of crowds. Subjective evaluations are mostly not consistent. Pairwise comparisons can support the exploration of inconsistences. Paired comparisons can also be initialized if ranked evaluations are available e.g. scores from 1-to-n about certain phenomena. The persons in particular are always consistent, but the population (the average person) can produce a lot of inconsistencies. Based on reports (without graph-analyses) it is also possible to generate a multidimensional index set about potential anomalies – but specific program codes are always necessary. Population can be divided according to sociological dimensions, thus inconsistences can also be derived for each group of a population. This makes it possible to explore potential differences in the (standard and/or scientific) human thinking.

Keywords: context free, GPS, automation, online engine, value survey

Introduction

Paired comparison of objects (e.g. phenomena, terms, keywords, teams, persons, countries, brands, etc.) seems to be a simple problem for the question: which objects are more important than others? The most trivial solution is: the more preferred an object is, the more important it is. Unfortunately, there are specific circumstances like the Simpson paradox[1] (where the winner becomes a loser day by day depending on the amount of daily scores). In addition, there can be specific inconsistent constellations, where e.g. A>B, B>C, but C>A! Pairwise comparisons can also be incomplete: independent islands of objects can be explored. If object (i) can be equal to object (j), then the evaluation of ranks of objects leads to newer problems like islands of identical objects and their connection with other lonely objects or identical object islands. Therefore paired comparisons can produce different types of anomalies with different volumes. A multidimensional evaluation tries to aggregate unique index values about anomalies and it tries to express the importance of objects based on only a single scale.

Evaluation ranking solutions should always be extremely robust: each possibility of subjectivity should be avoided. Human preferences may not play any role in them, especially using subjective scores for diverse dimensions of anomalies from pairwise comparisons is not preferred. It is necessary to involve anti-discriminative ranking methods (e.g. similarity analysis) in order to ensure that the scoring of dimensions and even their stairs will be derived from an optimized process.

The above mentioned problem complex is a part of the limited set of quasi GPS-solutions (general problem solvers), where context free solutions are offered for a wide spectrum of contextual similar problems (like building of object-chains according their importance). If a problem could be solved by logic, it should be automated and online as far as possible. Parallel solutions are always welcome – in

[1] https://de.wikipedia.org/wiki/Simpson-Paradoxon

order to derive the best solution (c.f. Occam's razor[2]). The (online) automation of solutions and the capability of choosing the best one is that, what Industry 4.0 really means for the scientific work! Motto: „Science is what we understand well enough to explain to a computer. Art is everything else we do." (Donald Ervin Knuth)

This article is the fifth item of a paper-series (see EDEN 2017, HASSACC 2017: articles about decision support mechanisms in learning management systems and pairwise comparison-based, report-driven algorithms for deriving index values for the following questions: How should an object ranking be built based on anti-discriminative models? Are a set of paired comparisons really different compared to random patterns?)

In this article, 13 evaluation dimensions about anomalies in pairwise comparison will be presented (only pivot reports were used instead of the original specific source codes like graph-approaches). The partial solutions (like pivots in Excel) already existing may be used later on a robust SQL platform, which makes scaling easier in the future in case of quasi unlimited volume of paired comparisons (e.g. data from sport events: searching for importance of teams/persons based on comparison in time and/or space – e.g. what is the best team/person? A similar (wide spectral) field of use would be the evaluation of answers from questionnaires.) The pairwise comparison-based evaluation is a typical context free problem: questionnaires about partial comparisons could be defined in education, enterprises, markets, etc.

This documentation serves the preparation of programming work steps (see ertekkutatas2v2.xlsx).

Problem identifications and their solutions

The chapters about elementary questions (hypotheses, anomalies) ensure a certain understanding of the anomalies' parallel layers of the originating in the characteristics of pairwise comparison. The index buildings requires randomized data patterns like Figure 1:

				atypical	53			changed relations	
			rel = =INT(VÉL()*4)+1	standard	82				
person_	object1	object2	relation_id	date	O2<O1	O1_s	O2_s	relation_id	o1o2
1	0	1	3	...	standard	0	1	3	01
1	1	1	2	...	standard	1	1	2	11
1	2	1	4	...	atypical	1	2	4	12
1	3	1	2	...	atypical	1	3	2	13
1	4	1	3	...	atypical	1	4	3	14
1	5	1	2	...	atypical	1	5	2	15
1	6	1	3	...	atypical	1	6	3	16
1	7	1	4	...	atypical	1	7	4	17
1	8	1	2	...	atypical	1	8	1	18
1	9	1	3	...	atypical	1	9	3	19
1	0	2	2	...	standard	0	2	2	02
1	1	2	3	...	standard	1	2	3	12
1	2	2	4	...	standard	2	2	4	22
1	3	2	2	...	atypical	2	3	1	23
1	4	2	2	...	atypical	2	4	1	24
1	5	2	2	...	atypical	2	5	2	25
1	6	2	2	...	atypical	2	6	1	26
1	7	2	2	...	atypical	2	7	1	27
1	8	2	3	...	atypical	2	8	3	28
1	9	2	3	...	atypical	2	9	3	29
1	0	3	3	...	standard	0	3	3	03

Figure 1: Randomized data (source: own presentation, where atypical pairs are standardized)

Index Nr.1. - Partiality

Phenomenon (relative): Ratio of partiality (ratio of lacks of positions) compared to the amount of objects, where the higher the rate of partiality is, the higher the personal consistence index could be.

[2] http://miau.gau.hu/miau/185/occams_razor_finetuned.doc

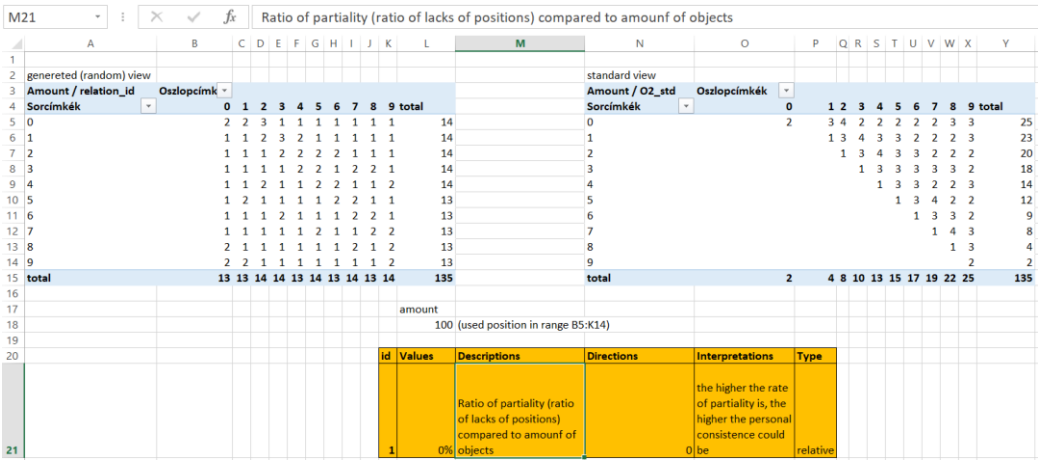
Question

How can the ratio be derived of partiality from pairwise comparison if 4 relation ids will be used?

- id=1: importance of object1 > importance of object2,
- id=2: importance of object1 < importance of object2,
- id=3: importance of object1 = importance of object2,
- id=4: importance of object1 * importance of object2, where the "*" stands for lack of decision.

Solution

A report-based solution starts from pivot-tables:



M21 — Ratio of partiality (ratio of lacks of positions) compared to amounf of objects

genereted (random) view

Amount / relation_id	Oszlopcímk									
Sorcímkék	0	1	2	3	4	5	6	7	8	9 total
0	2	2	3	1	1	1	1	1	1	14
1	1	1	2	3	2	1	1	1	1	14
2	1	1	1	2	2	2	1	1	1	14
3	1	1	1	1	2	2	1	2	2	14
4	1	1	2	1	1	2	2	1	1	14
5	1	2	1	1	1	1	2	2	1	13
6	1	1	1	2	1	1	1	2	2	13
7	1	1	1	1	1	2	1	1	2	13
8	2	1	1	1	1	1	1	2	1	13
9	2	2	1	1	1	1	1	1	2	13
total	13	13	14	14	13	14	13	14	14	135

amount: 100 (used position in range B5:K14)

id	Values	Descriptions	Directions	Interpretations	Type
1	0%	Ratio of partiality (ratio of lacks of positions) compared to amounf of objects	0	the higher the rate of partiality is, the higher the personal consistence could be	relative

standard view

Amount / O2_std	Oszlopcímkék									
Sorcímkék	0	1	2	3	4	5	6	7	8	9 total
0	2	3	4	2	2	2	2	2	3	3 25
1		1	3	4	3	3	2	2	2	3 23
2			1	3	4	3	3	2	2	2 20
3				1	3	3	3	3	3	2 18
4					1	3	3	2	2	3 14
5						1	3	4	2	2 12
6							1	3	3	2 9
7								1	4	3 8
8									1	3 4
9										2 2
total	2	4	8	10	13	15	17	19	22	25 135

Figure 2: Index Nr.1. - Partiality (source: own presentation, where 0% = there is no lack in the matrix on the left side, or no lack in the triangular view on the right side compared to 100%=10*10 ← amount of objects = 10)

The pairwise comparison can use pairs like A?B or B?A. Standardized view is, where the object ids will be transformed so, that the left side of a pair is always less than the right side (see triangular matrix above).

The non-standardized view makes possible to check, whether each position in the matrix is covered (see Figure 2). The amount of decisions for a given pair from object ids is not relevant in this phase.

Remark: the cross-tab-report (see Figure 2) can be supported in SQL in different ways.

Index Nr.2. – Multiple answers

Phenomenon (relative): Ratio of inconsistence in case of multiple asking for the same object pairs compared to total amount of decisions, where the less the amount of inconsistence in case of multiple answering is, the higher the personal consistence index could be.

Question

How can the amount of inconsistence be derived in case of multiple asking for the same object pairs compared to total amount of decisions?

Solution

Through reports like Figure 3, the average relation id (only id=1 and id=2) can be visualized and also the amount of records behind the average building process. Based on further IF/THEN constructions, the affected pairs can be identified. The index Nr.2. value (for multiply answers) will be calculated from the amount of the affected pairs and the amount of objects^2.

relation_id	(Több tétel)		<--only id=1 and id=2		
O1	O2	avg(relations)	amount(records)	amount	average
		Átlag / relation_id	Mennyiség / relation_id	more, than 1	not integer
0	0	2.00	1		
	1	2.00	1		
	2	1.50	2	1.50	1.50
	3	2.00	1		
	5	1.00	1		
	6	2.00	1		
	8	1.50	2	1.50	1.50
	9	1.50	2	1.50	1.50
1	1	2.00	1		
	3	1.50	2	1.50	1.50
	5	1.33	3	1.33	1.33
	6	2.00	1		
	8	1.50	2	1.50	1.50
2	3	1.67	3	1.67	1.67
	4	1.50	2	1.50	1.50
	5	1.67	3	1.67	1.67
	6	1.67	3	1.67	1.67
	7	1.00	1		
3	4	1.50	2	1.50	1.50
	5	1.50	2	1.50	1.50
	6	2.00	2	2.00	

Potential values	Amount of non-integer averages
1.33	2
1.50	16
1.67	3
Total	21

id	Values	Descriptions	Directions	Interpretations	Type
	21	Amount of inconsistence in case of multiple asking for the same object pairs	1	the less the amount of inconsistence is, the higher the personal consistence is	absolute
2	21%	ratio of inconsistence in case of multiple asking for the same pairing, where 100% means, each existing pairs in the pattern	1	the less the ratio of inconsistence is, the higher the personal consistence is	relative

Figure 3: Filtering multiply answers with inconsistences (source: own presentation, where the report on the left side is just a short abstract from the whole table with O1_id=max.8 and O2_id=max.9 – and 21 %=21/(10*10) ← amount of objects = 10).

Remark: The report for object1*object2 (see header of rows in Figure 3) with two fields (like average and amount of averages of relation ids 1 and 2) is a standard action in SQL. The further filtering needs specific program codes. But the calculation of the needed amount about not integer averages is also a standard SQL action.

Index Nr.3. – Chaos potential I.

Phenomenon (relative): Chaos potential I (it means that the amount of average=1.5 positions for relation ids 1 and 2) compared to the amount of potential positions (amount of object^2), where the less the chaos potential I is, the higher the personal consistence index could be.

Question

How can the amount of position be calculated having the value 1.5 as average of object ids for the given object pair variations?

Solution

From Figure 3 (column = average not integer), the amount of positions with the value 1.5 can be derived.

Potential values	Amount / averages of affected pair variations
1.33	2
1.50	16
1.67	3

id	Values	Descriptions	Directions	Interpretations	Type
	3	16%	chaos potential (it means: amount of average=1.5 position for relation ids 1 and 2) compared to the amount of potential positions	1	the less the chaos potential is, the higher the personal consistence is	relative

Figure 4: Chaos potential I (source: own presentation, where 16 %=16/(10*10) ← amount of objects = 10)

Remark: The amount of the benchmark of 1.5 can be derived through a standard SQL-action.

Index Nr.4. – Chaos potential II.

Phenomenon (relative): Chaos potential II (it means: average of absolute differences to average=1.5 position) compared to 1.5, where the less the chaos potential II is, the higher the personal consistence index could be

Question

How can be calculated the average of differences to the value 1.5?

Solution

The differences calculated based on the benchmark of 1.5 and the set of affected position can be transferred into a report (s. Figure 5).

Differences to benchmark = 1.5	Amount of positions	Average
0.167	5	0.167
Total and average	**5**	**0.167**

id	Values	Descriptions	Directions	Interpretations	Type
4	11%	chaos potential II. (it means: average of absolute differences to average=1.5 positions) compared to 1.5	1	the less the chaos potential is, the higher the personal consistence is	relative

Figure 5: Chaos potential II (source: own presentation, where 11 % = 0.167/1.5)

Remark: The amount and the average of differences to the benchmark of 1.5 can be derived through a standard SQL-action. But the filtering before (in order to build the differences) needs a special part in the program code.

Index Nr.5. – Lack of opinions

Phenomenon (relative): Ratio of the relation-id "4" compared to amount of total records, where the higher the ratio of the non-evaluation-id is, the higher the personal consistence index could be

Discussion: The direction "*the higher the ratio of the non-evaluation-id is, the higher the personal consistence index could be*" could also be formulated with an inverse logic: the less the ratio of the hidden force fields is, the higher the personal consistence index could be, because the lack of knowledge (the lack of decisions, opinions) may also be interpreted as a kind of risk to generate inconsistence (c.f. the more lack, the more inconsistence) or even as a kind of stability to not generate inconsistence through avoiding clear statements (c.f. the more lack, the less clarity = the less chance to explore inconsistences).

Question

How can the ratio be derived for the relation id "4"?

Solution

The relation id=4 means, the evaluator cannot, or does not want to express the relation between two objects.

		id	Values	Descriptions	Directions	Interpretations	Type
Relation ids	Amount / relation_id	5	24%	Ratio of the relation_id "4" compared to amount of total records	0	the higher the ratio of the non-evaluation-id is, the higher the personal consistence could be	relative
1	33						
2	35						
3	35						
4	32						
Total	**135**						

Figure 6: Hidden opinions (source: own presentation, where 24 %=32/135)

Remark: The report (Figure 6) for relation ids and their frequencies is a standard SQL-action.

Index Nr.6. – Sameness index

Phenomenon (relative): Ratio of the valid test positions compared of the total amount of the test position, where test position are the pairs with identical object ids – the higher the ratio of the valid test positions is, the higher the personal consistence index could be

Question

How can be derived the ratio of the valid test positions?

Solution

The attributes “same” and “different” for the object ids of an object pair can be defined in advance as a kind of default status variable. This attribute values can be used for the reporting (see Figure 7).

Amount / same_ids	Relation ids				
Object ids	1	2	3	4	Total
different	8	9	24	18	59
same	2	2	4	4	12
Total	10	11	28	22	71

id	Values	Descriptions	Dir	Interpretations	Type
6	33%	Ratio of valid test evaluations compared of the total amount of test positions	0	the higher the ratio of the valid evaluations is, the higher the personal consistence could be	relative

relation_id	(Több tétel)										(id=1, id=2)
Amount / relation_id	object ids										
objects ids	0	1	3	5	6	7	8	9	total		
0	1	1	1	1	1				5		
1	1				1				2		
2						1			1		
3					2	1		1	4		
4				1		1			2		
5								1	1		
6						1	1		2		
7								1	1		
8								1	1		
9								2	2		
total	1	2	1	2	4	4	1	6	21		
											check
											21
											check
											4

relation_id	(Több tétel)									
Average / relation_id	bject ids									
object ids	0	1	3	5	6	7	8	9		
0	2	2	2	1	2					
1		2			2					
2						1				
3					2	1		1		
4				2		1				
5								1		
6						1	2			
7								2		
8								1		
9								1		

Figure 7: Valid test positions (source: own presentation, where 33 % = 4/12)

Remark: The building of the status variable and their options belong to the definition of the database storing relation ids for pairs. The value for the status variable should always be calculated at once if a new pair and the connected relation id will be stored. The prompt calculation and the report about the status variable are standard SQL actions (see Figure 7 – reports on the bottom).

Index Nr.7. – STD-DEV

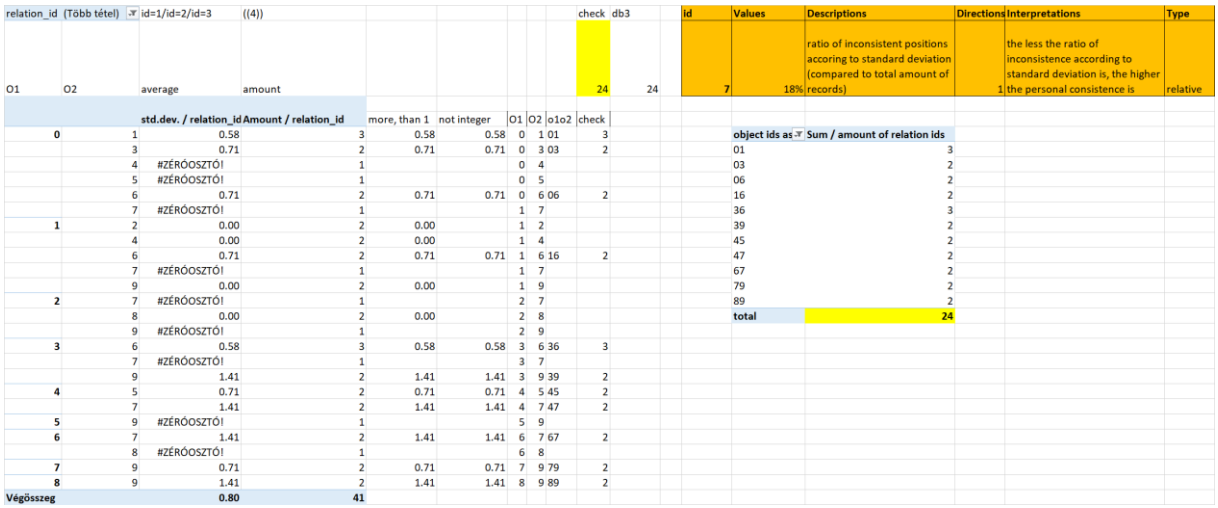
Phenomenon (…): Ratio of inconsistent positions according to the average of standard deviations of relation ids compared to the total amount of records), where the less the ratio of the inconsistence based on standard deviation is, the higher the personal consistence index is

Question

How can be derived the ratio of records affected by illogical average built from standard deviation values compared to the total amount of records?

Solution

A kind of inconsistence can be defined, if the amount of affected records will be compared to the whole amount of pairs in case of specific averages of standard deviations according to relation ids:



relation_id	(Több tétel)	id=1/id=2/id=3	((4))						check	db3
O1	O2	average	amount						24	24
		std.dev. / relation_id	Amount / relation_id	more, than 1	not integer	O1	O2	o1o2	check	
0	1	0.58	3	0.58	0.58	0	1	01	3	
	3	0.71	2	0.71	0.71	0	3	03	2	
	4	#ZÉRÓOSZTÓ!	1			0	4			
	5	#ZÉRÓOSZTÓ!	1			0	5			
	6	0.71	2	0.71	0.71	0	6	06	2	
	7	#ZÉRÓOSZTÓ!	1			1	7			
1	2	0.00	2	0.00		1	2			
	4	0.00	2	0.00		1	4			
	6	0.71	2	0.71	0.71	1	6	16	2	
	7	#ZÉRÓOSZTÓ!	1			1	7			
	9	0.00	2	0.00		1	9			
2	7	#ZÉRÓOSZTÓ!	1			2	7			
	8	0.00	2	0.00		2	8			
	9	#ZÉRÓOSZTÓ!	1			2	9			
3	6	0.58	3	0.58	0.58	3	6	36	3	
	7	#ZÉRÓOSZTÓ!	1			3	7			
	9	1.41	2	1.41	1.41	3	9	39	2	
4	5	0.71	2	0.71	0.71	4	5	45	2	
	7	1.41	2	1.41	1.41	4	7	47	2	
5	9	#ZÉRÓOSZTÓ!	1			5	9			
6	7	1.41	2	1.41	1.41	6	7	67	2	
	8	#ZÉRÓOSZTÓ!	1			6	8			
7	9	0.71	2	0.71	0.71	7	9	79	2	
8	9	1.41	2	1.41	1.41	8	9	89	2	
Végösszeg		0.80	41							

id	Values	Descriptions	Directions	Interpretations	Type
7	18%	ratio of inconsistent positions accoring to standard deviation (compared to total amount of records)	1	the less the ratio of inconsistence according to standard deviation is, the higher the personal consistence is	relative

object ids as	Sum / amount of relation ids
01	3
03	2
06	2
16	2
36	3
39	2
45	2
47	2
67	2
79	2
89	2
total	24

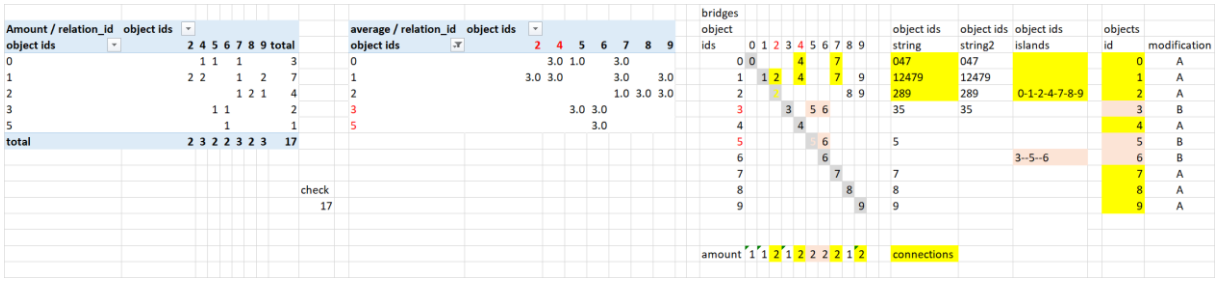
Figure 8: Inconsistence through standard deviation of relation ids (source: own presentation, where 18%=24/135)

Remark: The standard calculation of standard deviations for relation ids (without the id “4”) can be filtered by using the following rules: pairs (records), where only 1 record information is available, should not be involved into the result and also the records should not be involved, where the standard deviation of the relation ids (by multiply availability) exactly zero is. The “SELECT-ion” with filtering effects (“WHERE”) a standard SQL-action is.

Index Nr.8. – Inconsistent object islands

Phenomenon (relative): Ratio of inconsistences between the same islands (compared to the total number of records), where the less the ratio of the inconsistence between the same islands is, the higher the personal consistence index could be

The term “object island” means: Which object ids build an island (a set of objects) chained through the relation id “3” (describing the identity of object ids)?



Amount / relation_id	object ids								
object ids	2	4	5	6	7	8	9	total
0		1	1		1			3
1	2	2			1		2	7
2				1	2	1		4
3			1	1				2
5				1				1
total	2	3	2	2	3	2	3	17

check
17

average / relation_id	object ids						
object ids	2	4	5	6	7	8	9
0			3.0	1.0		3.0	
1	3.0		3.0			3.0	3.0
2						1.0	3.0	3.0
3					3.0	3.0	
5						3.0	

bridges object ids	0	1	2	3	4	5	6	7	8	9
0	0				4			7		
1		1	2		4			7		9
2			2						8	9
3				3		5	6			
4					4					
5						6				
6						6				
7							7			
8								8		
9									9	
amount	1	1	2	1	2	2	2	2	1	2

object ids	object ids	object ids	objects	
string	string2	islands	id	modification
047	047		0	A
12479	12479		1	A
289	289	0-1-2-4-7-8-9	2	A
35	35		3	B
			4	A
5			5	B
		3--5--6	6	B
7			7	A
8			8	A
9			9	A
connections				

Figure 9: Deriving islands based on reports (source: own presentation)

The primer database (db_rnd: in this case with 135 records) got filtered step by step: like

- db2: deleting records with inconsistent relation ids, where diverse relation ids are available for the same object-pairs
- db3: deleting records with inconsistent relation ids, where the same objects build pairs
- db4: deleting records, where relation id = 4 AND “storno” sign from db3 is given

After these reduction steps there are only object pairs, where the relation ids (in case of more than one record) are always the same (id=1, or id=2, or id=3) – see Figure 9 (second report from left to right).

Remark: These report are standard SQL-actions.

Islands of object ids can be detected, if bridges between object ids can be identified. A bridge is always a relation id “3”. If a bridge is existing between two object ids, then an island is born. A given object id can have more than one bridges. Therefore islands can have more than two object ids. The derivation of islands needs specific program codes where it is necessary to create a matrix Header for rows and columns are the object ids. Diagonal positions are identical with the affected object id. If the values in the report about averages of relation ids is “3”, then the position will be equal the object id from the column header. If the column header are higher in a column more than one, then the affected object id is a bridged object id. Affected objects id are the objects id from the row header. The chained object ids should be substituted with island ids (in this case through letters, like ABC).

Remark: The derivation of island ids needs a specific program code.

Potential constellations after creation of islands:

- It is possible that each object id belongs to the same island – it means: there is only one island – and therefore it is not possible to rank objects (each object is has the same importance).
- It is also possible, that there are no bridges – it means: each object id is independent – there are no islands.
- The most frequently situation is: there are just a few islands (where islands can have only one object id).

Question

How can be derived the ratio of inconsistent object islands?

Solution

After the substitution of numeric object ids with letters (island ids) for both members in an object pair, the islands ids can be involved into the reporting (see Figure 10).

relation_id	(Több tétel)		(id=1&id=2)
Amount / relation_id	**Islands**		
Islands	**A**	**B**	**Total**
A	1	2	3
B	2	1	3
Total	**3**	**3**	**6**

relation_id	(Több tétel)		(id=1&id=2)	
Average / relation_i	**Islands**			
Islands	**A**	**B**	**Total**	unit
A	1	1.5	1.33	1
B	1	1	1	
Total	**1**	**1.33**	**1.17**	

id	Values	Descriptions	Directions	Interpretations	Type
8	1.5%	Ratio of inconsistences between the same islands (compared to the total number of records)	1	the less the ratio of the inconsistence between the same islands is, the higher the personal consistence index could be	relative

Figure 10: Ratio of inconsistent islands (source: own presentation, where 1.5 %=(1+1)/135)

Remark: Filtered report, where only the relation ids “1” and “2” are visible, can be executed as a standard SQL-action.

Index Nr.9. – Rationality of islands

Phenomenon (relative): Ratio of rational relations between different islands (compared to the total number of records), where the higher the ratio of the rational relations between different islands is, the higher the personal consistence could be.

Question

How can the ratio be derived of rational relation between different islands?

Solution

The rationality of islands as variables for anomaly description means that the relation between islands like A?B and B?A must be consequent. If A>B, then B<B and vice versa.

relation_id	(Több tétel)		(id=1&id=2)
Amount / relation_id	**Islands**		
Islands	**A**	**B**	**Total**
A	1	2	3
B	2	1	3
Total	**3**	**3**	**6**

relation_id	(Több tétel)		(id=1&id=2)
Average / relation_	**Islands**		
Islands	**A**	**B**	**Total**
A	1	1.5	1.33
B	1	1	1
Total	**1**	**1.33**	**1.17**

id	**Values**	**Descriptions**	**Directions**	**Interpretations**	**Type**	unit
9	2.2%	Ratio of rational relations between different islands (compared to the total number of pairs)	0	the higher the ratio of the rational relations between different islands is, the higher the personal consistence could be	relative	1

Figure 11: Ratio of rational relations between different islands (source: own presentation, where the amount of the logical set relation is calculated – therefore: 2.2 % = 2+2-1/135 and “unit=1” describe the amount of irrational relation compared to the further relation ids - see Figure 12)

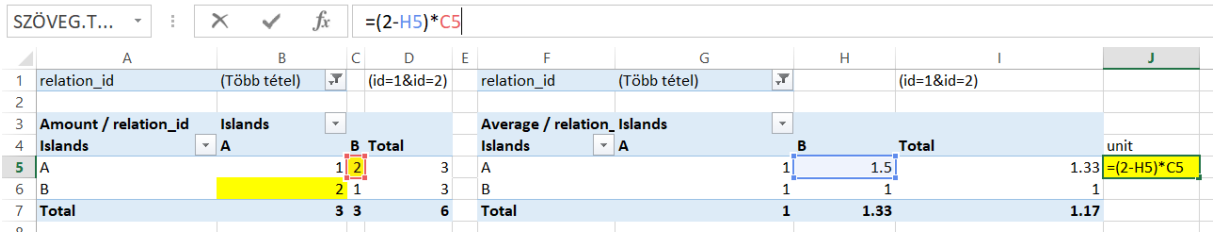


Figure 12: Calculation of the amount of irrational relation ids (source: own presentation, where the average of the relation ids set the sign for non-integer positions).

Remark: Calculation of a unit needs a specific program code in case of non-integer averages, but the reports before are standard SQL-actions.

Index Nr.10. – Rational islands

Phenomenon (relative): Ratio of island-pairs with the same rational relations for both islands compared to the potential amount of pairs of islands with different islands, where the less the ratio of the islands with lack of preferences is, the higher the personal consistence could be.

Question

How can the ratio of rational islands be derived?

Solution

Figure 10-11-12 show, that the amount of island ids is 2. Therefore the amount of the reported positions (island-pairs) is 4 (2*2). The diagonal positions are for the same island ids as a pair. If the island pair with different island ids are covered through experiences, then the lack of information is zero.

id	**Values**	**Descriptions**	**Directions**	**Interpretations**	**Type**
10	0%	Ratio of island-pairs with the same rational relations for both islands compared to the potential amount of pairs of islands with different islands	1	the less the ratio of the islands with lack of preferences is, the higher the personal consistence could be	relative

Figure 13: Ratio of lack of information concerning island pairs with different ids (source: own presentation, where 0 % = there is no lack of information in the example)

Remark: The reports for deriving lacks are standard SQL-action. The hermeneutics for the reports should be programmed through a specific code.

Index Nr.11. –Independent islands

Phenomenon (relative): Ratio of independent islands/objects compared to all islands objects, where the less the ratio of the independent objects is, the higher the personal consistence index could be

Independent objects/islands mean: there are no connections to the other objects/islands. About independencies can only be spoken, if the objects/islands cannot be ranked in a holistic way where each object/island id belongs to the evaluation chain (see Figure 14).

To build evaluation chains it is enough if only one relation id for each object/island id is available. The transitivity can be checked through a kind of relative score (see Figure 14 – on the bottom). The relative score can be derived, if a matrix will be defined, where the rows and the columns are the object/island ids. The diagonal positions have always a zero value. The triangular half is to mirror with converted relation ids (instead of “1” always “2” and vice versa). After these preparations, the sums of the rows and the columns will be calculated. The relative score will be calculated from the sums of rows and columns in the way that these position will be divided with the amount of the non-zero positions in the affected row or column.

The transitivity can also be simulated through leveled transformations (see Figure 14), but these steps are more complex than the process of the relative scores – therefore the relative scores are more robust and this way should be used to building ranks for objects/islands based on reports.

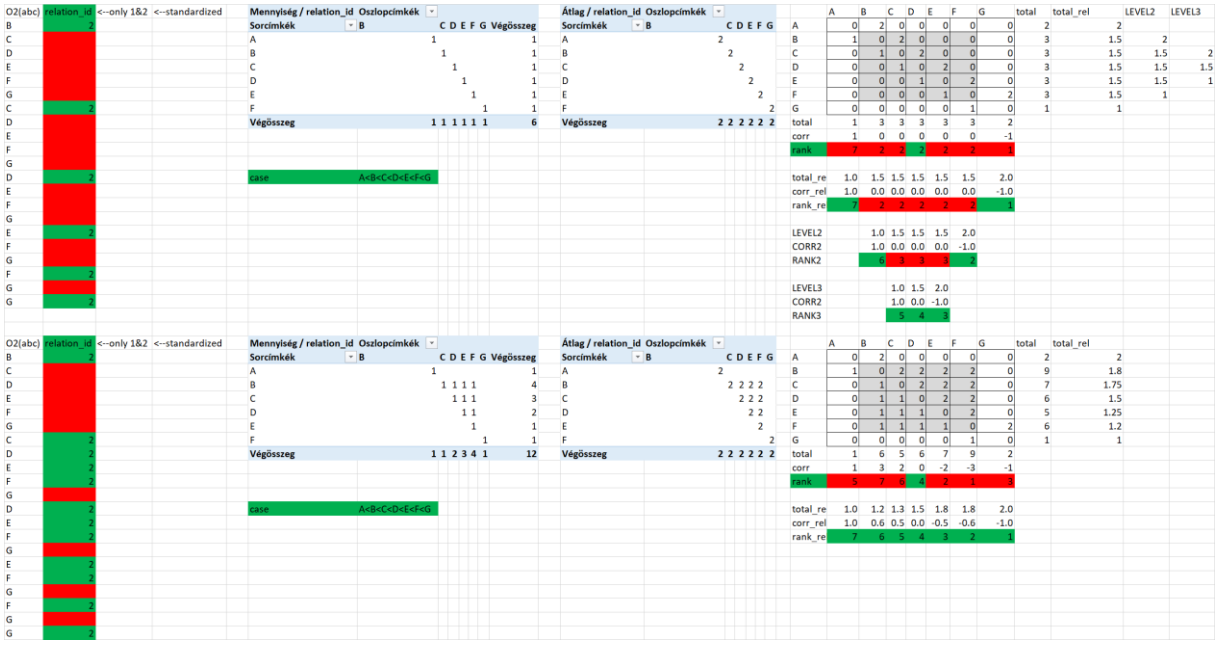


Figure 14: Demonstration of transitivity based on reports (source: own presentation)

Question

How can the ratio be derived of independent islands/objects?

Solution

The reports based on objects/islands ids for the row and column headers can be evaluated as follows: the lacks of ids/letters can be identified with a specific program code, and the sum of the covered ids/letters can be divided by the doubled amount of objects/islands.

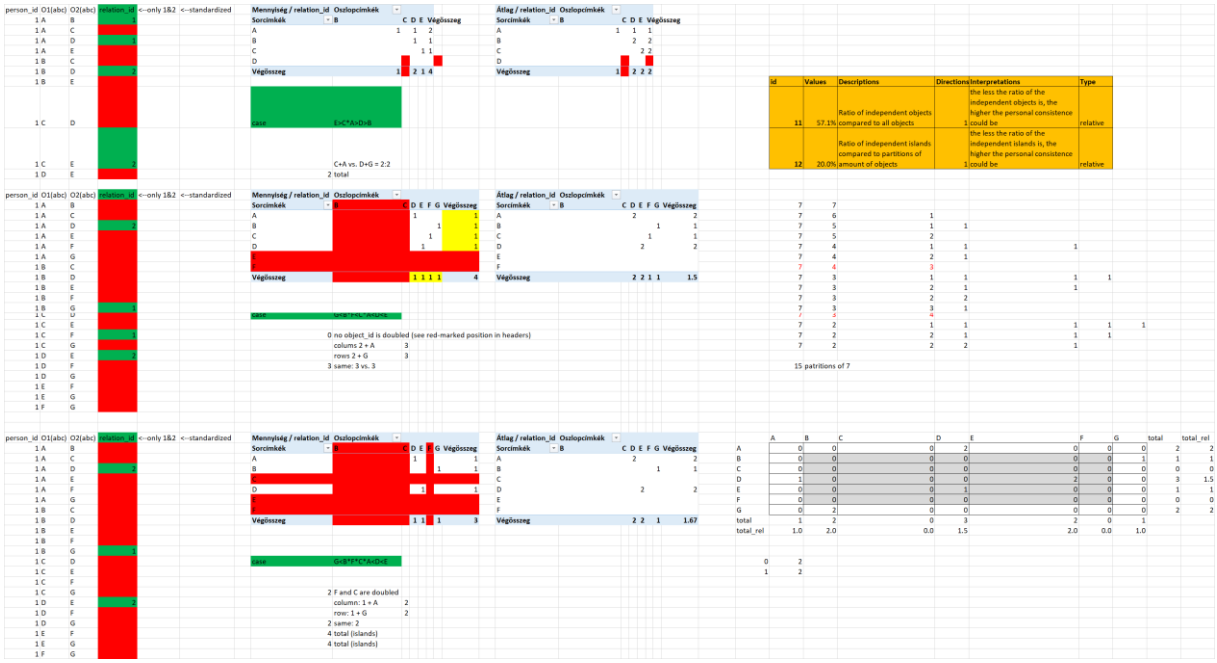


Figure 15: Demonstration of independent islands/objects (source: own presentation, where 57.1 % = ([1+1+1+1]+ [1+1+1+1])/(7+7) ← amount of objects/island = 7)

Remarks: The reports are standard SQL-actions. The evaluation of the reports needs a specific program code.

Index Nr.12. – Independent islands

Phenomenon (relative): Ratio of independent islands compared to partitions of amount of objects/islands, where the less the ratio of the independent islands is, the higher the personal consistence index could be

The amount of independent islands is a sign for inconsistence, because the independency is antagonistic to ranking.

Partitions: in number theory and combinatorics, a partition of a positive integer n, also called an integer partition, is a way of writing n as a sum of positive integers. Two sums that differ only in the order of their summands are considered the same partition. (If order matters, the sum becomes a composition.) About partitions in detail: http://tamop412a.ttk.pte.hu/files/Kombinatorika_kesz_jav3_final.pdf (S.37.) or https://en.wikipedia.org/wiki/Partition_(number_theory)

The amount of independent islands can be compared to the amount of partitions or to the amount of objects/islands.

Question

How can the ratio of independent islands be derived?

Solution

Figure 15 (on the right side / bottom) shows that the matrix used in the calculation scheme for relative scores produces specific relative score values in case of availability of inconsistencies:

- score = zero – it means: lonely objects/islands (amount of zero values is the amount of lonely objects/islands)
- score = 1 or score = 2: other islands (where the amount of score=1 is the same amount as for score=2 therefore the sum of these groups leads to the amount of other islands)
- score = other: irrelevant for the calculation

Remark: The calculation of the amount of independent islands needs the same program code as before in case of relative scores.

Index Nr.13. – Corrupted chains

Phenomenon (relative): Ratio of rational ranks compared to all objects, where the higher the ratio of the rational ranks is, the higher the personal consistence index could be

Question

How can the rational part of an evaluation chain be derived?

Solution

Corrupted chains means: there are one or more specific constellation like A>B, B>C, C>A instead of C<A. Corrupted chains can not be ranked from 1 to n, where n is the amount of the objects/islands.

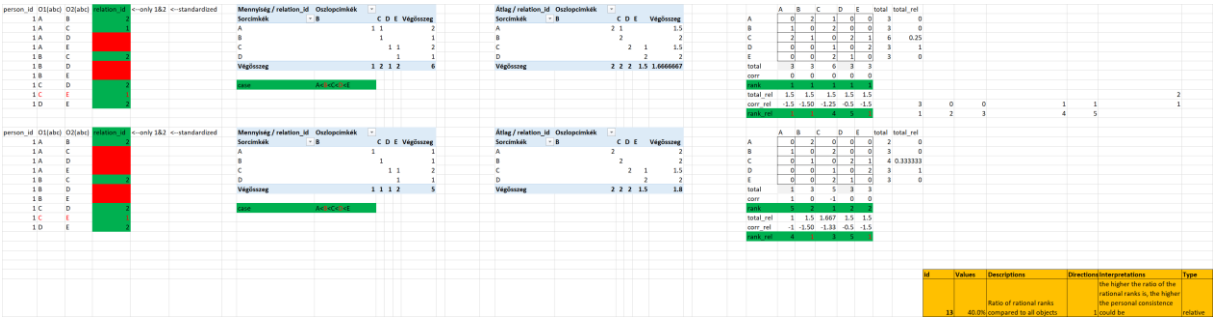


Figure 16: Corrupted chains (source: own presentation, where 40 % = 2/5, it means amount of non-repeated ranks divided by amount of objects/islands)

Remark: The building of relative scores leads to ranks as a kind of side-effect. The exploring of repetition of ranks needs a specific program code.

Index Nr.14.

Phenomenon (…): further attributes are possible without any limitation

Question

Each relevant question should lead to an attribute having direction (the-more-the-more or the-more-the-less).

Solution

Each question should be answered through algorithms…

Anti-discriminative models for personal consistence index

The above descripted anomaly indices deliver raw values for each person voting about pairs (about the importance of objects compared to each other). The anomaly indices has always a direction (see the more/the more, or the less/the more constructions before). Based on these directions, ranked indices can be derived. The ranked anomaly indices belong to an OAM (object-attribute-matrix), where the persons are the objects, and the attributes are the anomaly indices. Anti-discriminative models (direct and invers similarity analyses – see more: http://miau.gau.hu/myx-free/) are capable of creating a unique index from the multidimensional anomalies. This aggregated index has no subjective scores in the calculation scheme. Each stair-level in the staircase function comes from an optimizing approach.

The aggregated index is the personal consistence index for persons, who has a symmetric partial results after direct and invers similarity analysis. The lacks of symmetry are signs for non-evaluable persons. Person with symmetric results can be evaluated as consistent (above the norm), as inconsistent (below the norm) or norm-like.

Personal consistence index values can also be interpreted as a time series, if a person has more possibilities to vote about importance of objects.

Discussions

The steps presented above deliver a few indices about potential anomalies. Mostly, it is possible to use reports, but it seems to be unavoidable to derive personal consistence index without any specific program codes.

Conclusions

Other (e.g. graph-oriented) calculation models also exist. Parallel solutions can be compared to each other (see Occam's razor) if attributes for models will be defined in order to describe partially what seems to be good or wrong in the particular models. The multi-dimensional evaluation with anti-discriminative modelling is also valid for model evaluation (not only for aggregation of anomaly indices).

References

1. Pitlik, L. (2014): Occam borotvája finomhangolás után, Gödöllő, Magyar Internetes Alkalmazott/Agrárinformatikai Újság, No.185, 5p, 1-5p.
2. Simpson, E. H. (1951): The Interpretation of Interaction in Contingency Tables. In: Journal of the Royal Statistical Society, Ser. B. Band 13, S. 238–241.