

Tato prezentace je součástí wiki-prezentace [Metoda GUHA, LISp-Miner a typové úlohy](#)

Je dostupná z [této adresy](#)

Verze 12. 3. 2020

Typ úlohy: histogram roste x histogram klesá

Data: [Hotel](#)

Problém: *Existuje segment hostů a dvojice států tak, že pro tento segment jsou pro jednotlivé státy trendy spokojenosti opačné?*

Jan Rauch

Katedra informačního a znalostního inženýrství

Vysoká škola ekonomická v Praze

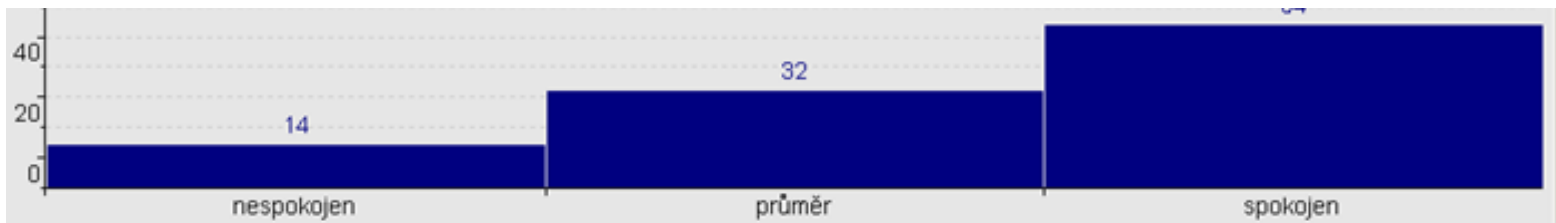
© Jan Rauch

Příklad analytické otázky

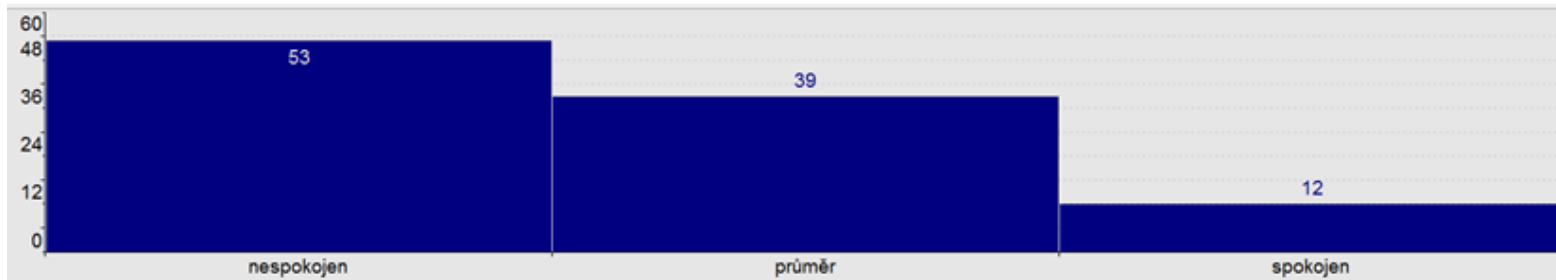
HOTEL:

Existuje segment hostů a dvojice států tak, že pro tento segment jsou pro jednotlivé státy trendy spokojenosti opačné? **Stát(?) x Stát(?) [DHodnocení ↑↓/?]**

Stát 1



Stát 2



Stát(?) x Stát(?) [DHodnocení ↑↓/?]

– příklad zadání procedury SDCF-Miner

ATTRIBUTES FOR HISTOGRAM

Attributes
» DHodnoceni(3 categories)

DHodnocení

Total length: -

QUANTIFIERS

Type	Range	Rel. Value	Units
SUM FirstSet	all	>= 20.00	Abs
S-UP FirstSet	all	>= 2.00	Abs
SUM SecondSet	all	>= 20.00	Abs
S-DN SecondSet	all	>= 2.00	Abs

Mode: Standard

SDCF-kvantifikátor

CONDITION

Pobyt Con, 1 - 4
 » PNoci_enum_m (seq), 1 - 2 B, pos
 » PDenTydne (subset), 1 - 1 B, pos
 » POsob (seq), 1 - 3 B, pos
 » POsobonoci_ef5 (seq), 1 - 2 B, pos
 Host Con, 0 - 2
 » HPohlavi (subset), 1 - 1 B, pos
 » HVek_exp (seq), 1 - 2 B, pos
 Meteo Con, 0 - 2
 » MObloha (subset), 1 - 1 B, pos
 » MTeploa_exp (seq), 1 - 2 B, pos

Relevantní segmenty hostů

(1) FIRST SET

Default Partial Cedent Con, 1 - 1
 » HStat (subset), 1 - 1 B, pos

1. stát

Total length: 1

(2) SECOND SET

Default Partial Cedent Con, 1 - 1
 » HStat#2 (subset), 1 - 1 B, pos

2. stát

Total length: 1

Jedno z mnoha možných zadání pro řešení dané analytické otázky

Total length: 1 - 8

Zadání SDCF-kvantifikátoru

List of SDCF quantifiers settings

Task: SDCF-Miner, opačné trendy spokojenosti (02)

Type	Source	Range	Rel.	Value	Units
Sum of frequencies	FirstSet	all	>=	20.00	Absolute number
Steps-up	FirstSet	all	>=	2.00	Absolute number
Sum of frequencies	SecondSet	all	>=	20.00	Absolute number
Steps-down	SecondSet	all	>=	2.00	Absolute number

- Segment hostů má ve státu 1 alespoň 20 členů
- Oba schody v histogramu pro stát 1 jsou nahoru
- Segment hostů má ve státu 2 alespoň 20 členů
- Oba schody v histogramu pro stát 2 jsou dolů

Example attribute: DHodnoceni (3 categories)

Category	Count
nespokojen	570
průměr	950
spokojen	480

0 475.00 950

Abs Rel sum Rel max

Buttons: Detail, Del, Up, Down, Copy, Paste, Add distribution quantifier, Add simple frequencies quantifier, Add step-up/down quantifier, Close

Stát(?) x Stát(?) [DHodnocení ↑↓/?]

– výstup procedury SDCF-Miner

Task run

Start: 28.3.2016 12:59:39 Total time: 0h 17m 42s

Number of verifications: 6633022

Number of hypotheses: 10 Mode: Standard

Add group Del group Edit group

Actual group of hypotheses: All hypotheses

Hypotheses in group: 10 Shown hypotheses: 10 Highlighted: 0

Delete hypotheses

Nr.	Id	DA-Sum	1:Sum	2:Sum	Hypothesis
1	10	85.000	157.00	72.00	DHodnoceni : HStat(ČR) × HStat(Německo) / POsob(2) & POsobonoci(>=vyšší)
2	9	36.000	69.00	37.00	DHodnoceni : HStat(ČR) × HStat(Rakousko) / POsob(1) & HPohlavi(žena) & HVek(>=od 28 do 60) & MTeplota(zima,neutrální)
3	3	31.000	53.00	64.00	DHodnoceni : HStat(Rakousko) × HStat(ČR) / PNoci(2) & PDenTydne(Pá) & POsob(<=2) & POsobonoci(nižší,průměr) & MObloha(slunečno)
4	1	31.000	53.00	64.00	DHodnoceni : HStat(Rakousko) × HStat(ČR) / PNoci(2) & PDenTydne(Pá) & POsob(<=2) & MObloha(slunečno) & MTeplota(zima,neutrální)
5	7	31.000	53.00	64.00	DHodnoceni : HStat(Rakousko) × HStat(ČR) / PNoci(2, <3;6>) & PDenTydne(Pá) & POsob(<=2) & POsobonoci(nižší,průměr) & MObloha(slunečno)
6	5	31.000	53.00	64.00	DHodnoceni : HStat(Rakousko) × HStat(ČR) / PNoci(2, <3;6>) & PDenTydne(Pá) & POsob(<=2) & MObloha(slunečno) & MTeplota(zima,neutrální)
7	4	27.000	43.00	54.00	DHodnoceni : HStat(Rakousko) × HStat(ČR) / PNoci(2) & PDenTydne(Pá) & POsob(<=2) & POsobonoci(nižší,průměr) & HVek(>=od 28 do 60)
8	2	27.000	43.00	54.00	DHodnoceni : HStat(Rakousko) × HStat(ČR) / PNoci(2) & PDenTydne(Pá) & POsob(<=2) & HVek(>=od 28 do 60) & MObloha(slunečno)
9	8	27.000	43.00	54.00	DHodnoceni : HStat(Rakousko) × HStat(ČR) / PNoci(2, <3;6>) & PDenTydne(Pá) & POsob(<=2) & POsobonoci(nižší,průměr) & HVek(>=od 28 do 60)
10	6	27.000	43.00	54.00	DHodnoceni : HStat(Rakousko) × HStat(ČR) / PNoci(2, <3;6>) & PDenTydne(Pá) & POsob(<=2) & HVek(>=od 28 do 60) & MObloha(slunečno)

Stát(?) x Stát(?) [DHodnocení ↑↓/?]

– ukázky z detailního výstupu procedury SDCF-Miner

Attribute: DHodnoceni
First set: HStat(ČR)
Second set: HStat(Německo)
Condition: POsob(2) & POsobonoci(vyšší, nejvyšší)

HStat(ČR)



HStat(Německo)

