

MTAT.03.264, Andmebaasid, 4. loeng
Relatsioonid ja
Relatsioonialgebra

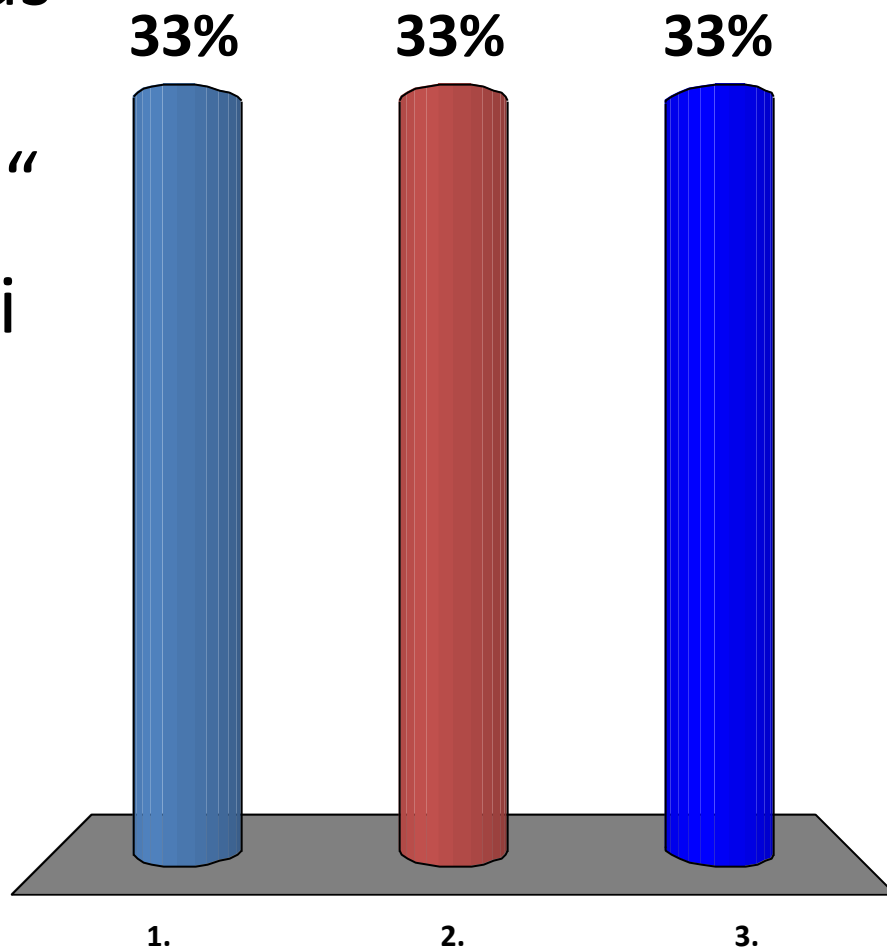
Anne Villems

7.03.2013

Sissejuhatuseks 2.loengu küsimusi

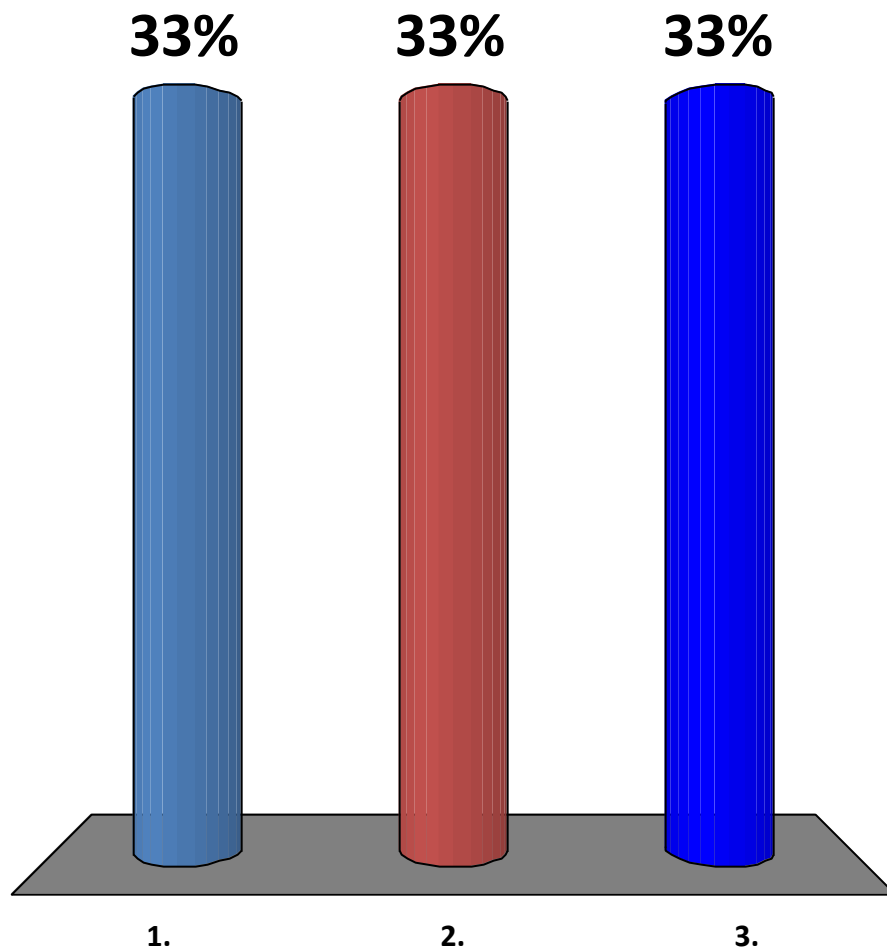
Inimeste nimed – kuidas modelleerida?

1. Üks tunnus „Nimi“, kus on nii ees- kui perenimi n. „Leo Lubi“
2. Kaks tunnust: Eesnimi ja Perenimi
3. Piisab eesnime esitähdest perenime ees: Nimi – „L.Lubi“



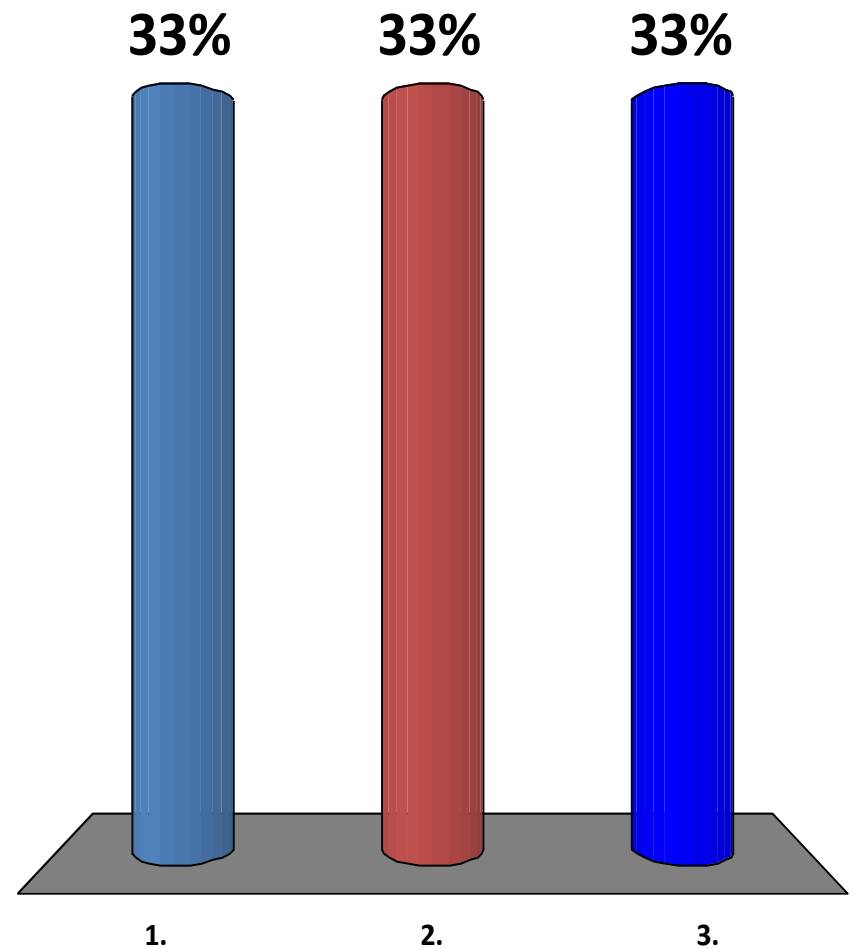
Vanus vs. Sünnikuupäev – kumb panna andmebaasi?

1. Vanus – hea teha päringuid vanuse järgi
2. Sünnikuupäev – on täpsem
3. Mõlemad – mugav kasutada seda, mida vaja



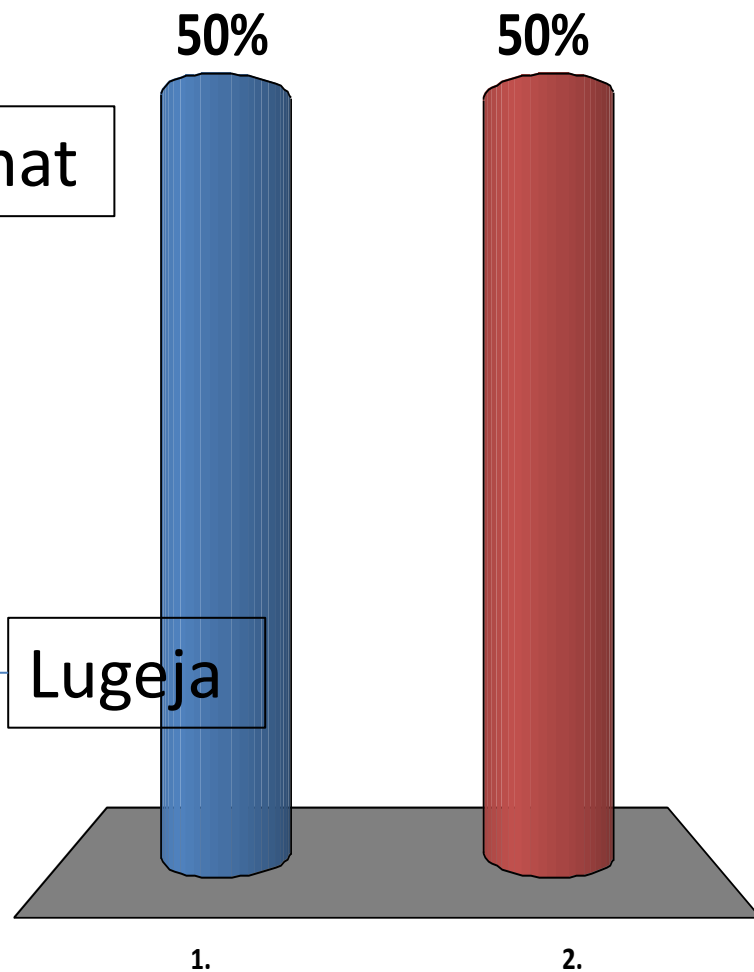
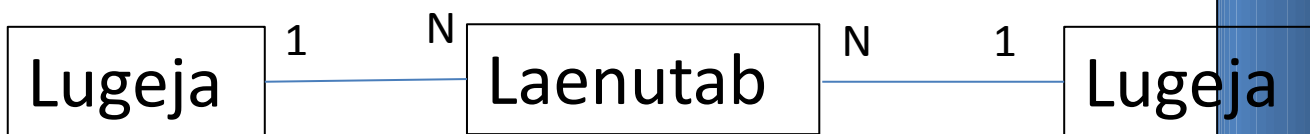
Aeg! Kas andmebaas käsitleb ainult antud ajahetke või pikemat perioodi?

1. Antud hetke, s.t. raamat – laenutaja on N:1 seos.
2. Pikemat perioodi, s.t. raamat-laenutaja on N:M seos.
3. Teha 2 baasi, hetkeolukord ja arhiiv



Kumb teile rohkem meeldib?

1. Kahe olemiga:



1) Relatsiooniline andmemudel

Relatsioonilise mudeli lühiülevaade (1)

- Olgu D_i lihtväärtuste hulk
 - Näiteks: eesnimede hulk, matriklinumbrite hulk jne.

- Relatsiooni definitsioon:

$$R \subseteq D_1 \times D_2 \times \dots \times D_n \quad (*)$$

- Hulgad D_i ei pea olema lõikumatud
- Näide: punktid 3-mõõtmelises ruumis $P(x,y,z)$
 - siin $D_1 = D_2 = D_3$, reaalarvude hulk
- Antud näites relatsioon on punktihulk. Näit.
 $\{ \langle 1,4,5 \rangle, \langle 2,7,1 \rangle, \langle 3,3,3 \rangle \}$

Relatsioonilise mudeli lühiülevaade (2)

- NB! Relatsioon on hulk s.t
 - korduvaid elemente pole
 - korteežide (kui hulga elementide) järjekord pole oluline
- Anname hulkade D_i esinemistele definitsioonis (*) nimed – siis ei pea iga korteeži elemendiga tema järjekorranumbrit kaasas vedama.
- Tähistame nimesid A_1, A_2, \dots, A_n
- Kui mingi D esineb relatsiooni kirjelduses mitu korda, on tema iga esinemine kandma erinevat nime.

Relatsioonilise mudeli lühiülevaade (3)

Nimede-alased kokkulepped siin kursusel:

- Mistahes sümbolite jada,
- Kui jutt konkreetsest relatsioonist, kasutame mõtestatud ja arusaadavaid nimesid
 - Näit. Eesnimi, eksami kuupäev, protokoll number
- Abstraktsel tasemel kasutame suurtähti alfabeedi esiootsast, vajadusel indeksitega
 - Näit: A, B, C või A_1, A_2, \dots, A_n või B_1, B_2, \dots, B_m

Relatsioonilise mudeli lühiülevaade (4)

Relatsiooni R kirjelduseks loeme kirjutist

$$R (A_1, A_2, \dots, A_n) \quad (**)$$

- Nimetame A_i relatsiooni atribuudiks (e. tunnuseks).
- Kuna atribuutide järjekord ei ole oluline, siis loeme kirjeldusega atribuutide järjekorra fikseerituks.
- Relatsiooni järk on tema atribuutide arv
- Relatsiooni (**) järk on n .

Relatsioonilise mudeli lühiülevaade (5)

- Relatsiooni kirjeldusele vastab hulk võimalikke relatsiooni seisundeid.
- Relatsiooni seisundit võib vaadata tabelina, mille veergude nimed on A_1, A_2, \dots, A_n , veergudes on väärtused vastavast hulgast D_i ja rida moodustab korteeži $\langle t_1, t_2, \dots, t_n \rangle$
- NB! Tabelis ei ole korduvaid ridu ja ridade jrk. ei ole oluline!

Relatsiooni näide

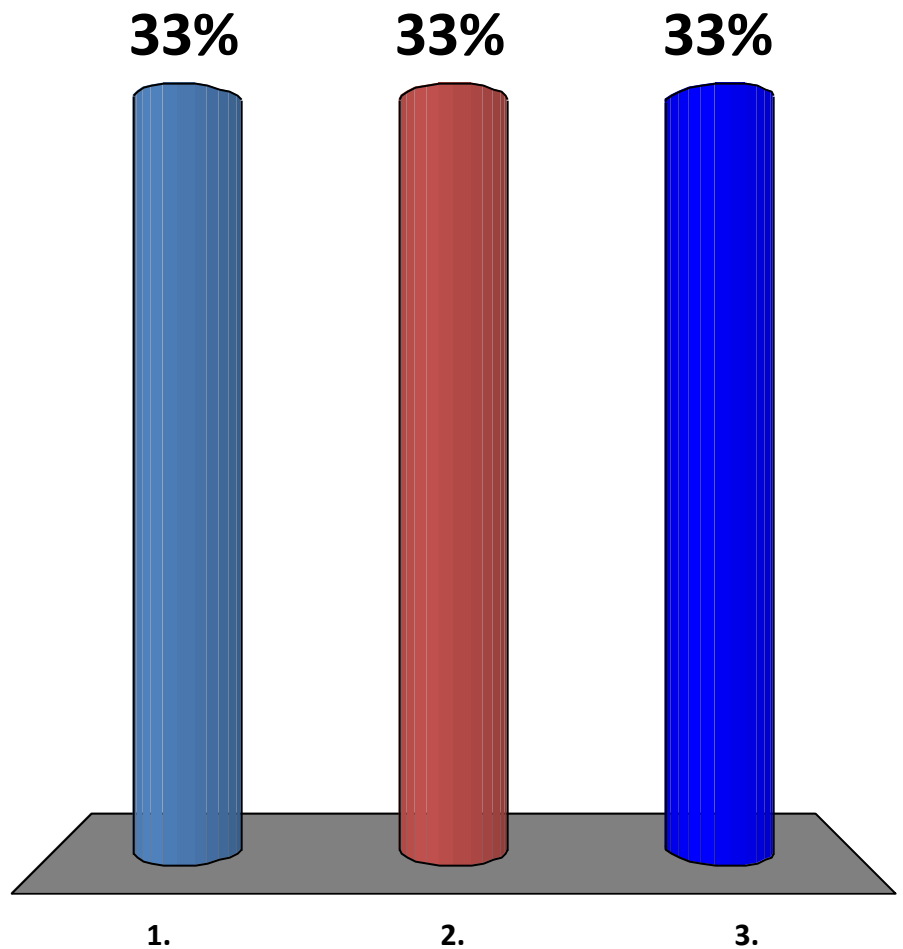
Matrikli number	Eesnimi	Perekonnanimi	Aine kood	Tulemus
A05555	Anu	Tark	MTAT.03.100	A
B16666	Ervin	Eriline	MTAT.03.191	arvestatud
A93218	Peeter	Python	MTAT.03.100	C
A74466	Maret	Muru	MTAT.03.100	B
B22277	Oskar	Orro	MTAT.03.100	B
B26888	Hannes	Haab	MTAT.03.142	arvestatud

Andmebaas relatsioonilises mudelis

- **Andmebaas** on relatsioonilise mudeli seisukohalt **relatsioonide hulk**.
- Andmebaasi kirjeldus on relatsioonide kirjelduste hulk.
- NB! Peale relatsioonide relatsioonilises mudelis muid vahendeid andmete esitamiseks ei ole!
- Nii olemid kui nendevahelised seosed - mõlemad esitatakse relatsioonide abil.

Kas relatsioon on maatriks?

1. Jah, sest ta on ju tabel
2. Ei, sest ta on hulk
3. Ei tea



Relatsioonialgebra

Tehted relatsioonidega

Üldised märkused relatsioonialgebra kohta

- Operatsioonid tehakse terve(te) relatsiooni(de)ga, mitte tema elementidega, s.t. mitte ridadega eraldi.
- Tegemist on algebraga s.t. relatsioonidest tehakse relatsioonid.
- Põhioperatsioonid on 6: kaks unaarset (argumendiks üks relatsioon) ja neli binaarset (argumendiks kaks operatsiooni).

Hulgateoreetilised operatsioonid

- Olgu R ja S relatsioonid: $R \cup S$, $R \cap S$, $R - S$
- **Kooskõlalisuse** nõue:
 - Relatsioonidel järk sama
 - Atribuutide vahel saab korraldada sellise $1 \leftrightarrow 1$ vastavuse, et vastavad andmehulgad D kattuvad
- Näiteks:

Tüdrukud (Eesnimi, IK, Perenimi, Pikkus, Kaal)

Poisid (Pnimi, Enimi, Isikukood, Raskus, Kasv)

A diagram consisting of four blue lines that cross each other. Two lines connect 'Eesnimi' to 'Pnimi' and 'Enimi'. Two other lines connect 'Perenimi' to 'Isikukood' and 'Pikkus' to 'Raskus'. The lines for 'Kaal' and 'Kasv' do not have corresponding connections.

Otsekorrutis

- $R \times S$ kus R järk on n ja S järk m
- Saab alati teha.
- Tulemusel on $n+m$ atribuuti;

R:

A1	A2	A3
a	2	3
b	2	3
c	1	1

S:

B1	B2
100	xx
200	yy

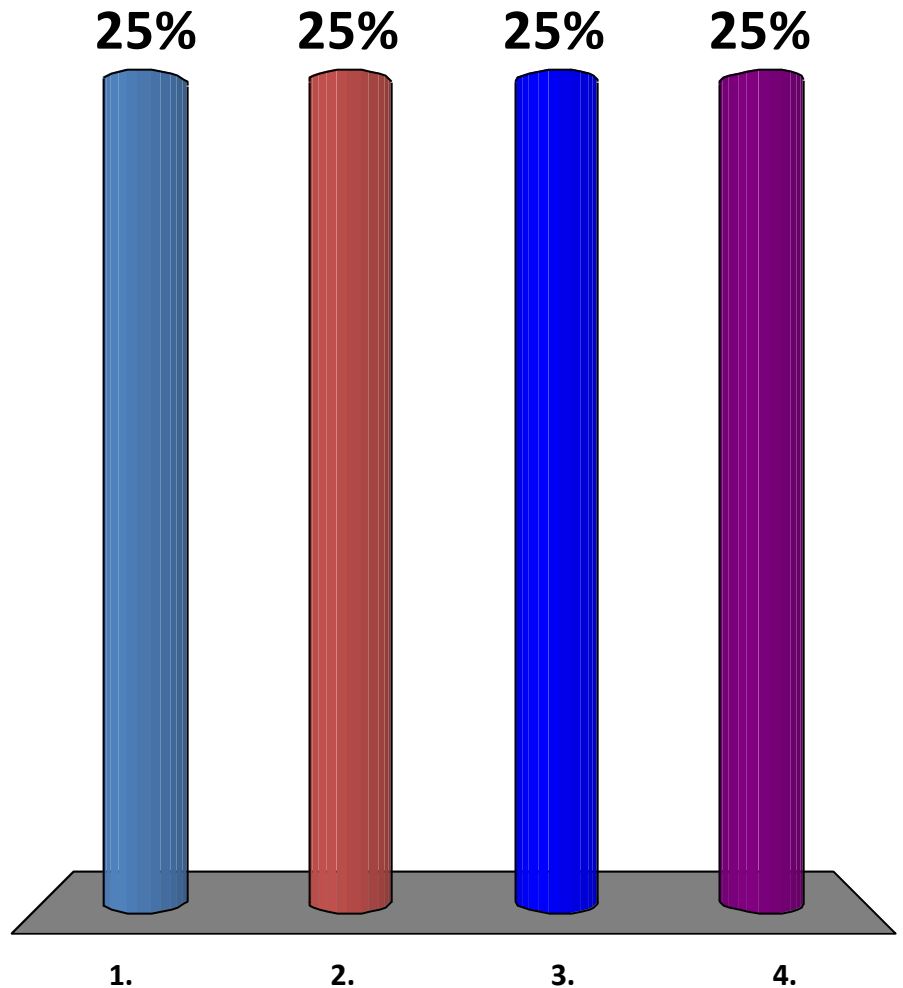
$R \times S$

A1	A2	A3	B1	B2
a	2	3	100	xx
a	2	3	200	yy
B	2	3	100	xx
B	2	3	200	yy
C	1	1	100	xx
c	1	1	200	yy

Kui R seisundil N rida
 S seisundil M rida
siis tulemusel $N \times M$ rida

Olgu R seisund tühi hulk ja S seisundis N rida. Siis Otsekorrutises $R \times S$ on:

1. N rida
2. 0 rida
3. 1 rida
4. Ei tea, mitu rida



Projektsioon

- $\pi_{i_1, i_2, \dots, i_k} (R)$
- Tulemuseks on relatsioon järku k , kus R seisundi igast korteežist jäetakse alles ainult valitud tunnused i_1, i_2, \dots, i_k .

- Näide:

$\pi_{A,B} (R)$

A	B
1	2
3	1

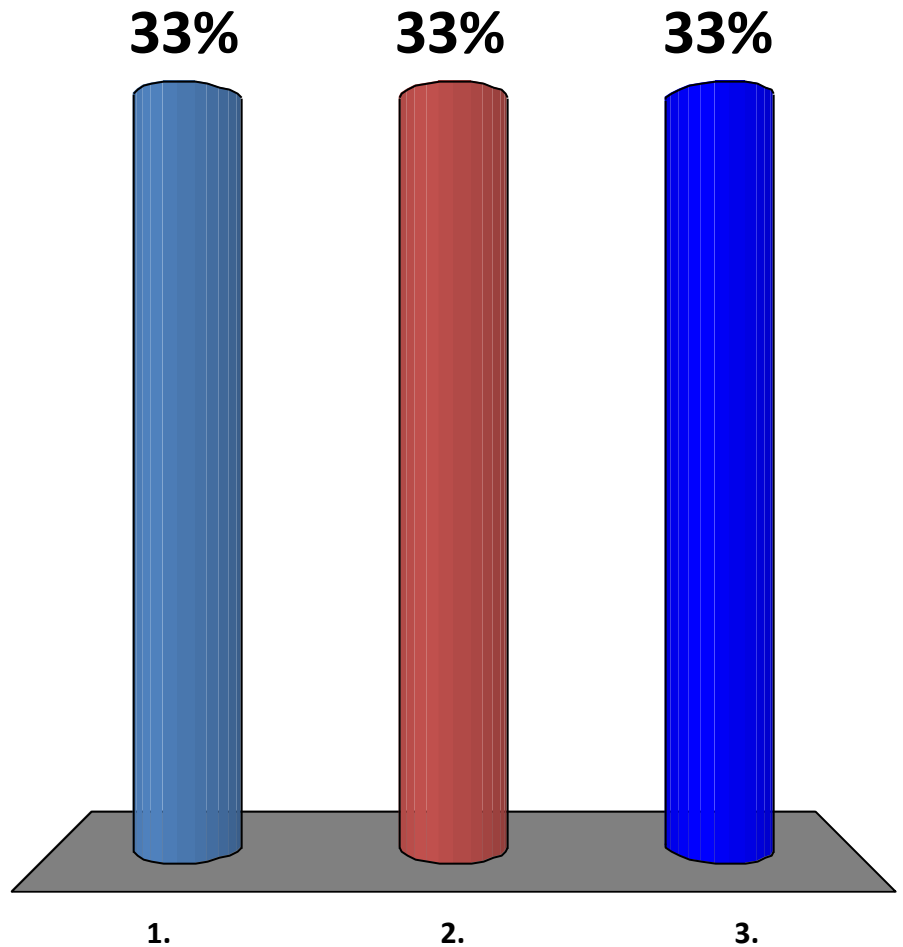
R

A	B	C
1	2	3
3	1	2
1	2	1

Kuhu kadus 3. rida?

Olgu relatsiooni R seisundis N rida. Kas projektsiooni tulemuseks saab olla tühi hulk?

1. Jah, saab küll.
2. Ei saa.
3. Ei tea.



Filter

- $\sigma_{\alpha} (R) = \{ \langle t_1, t_2, \dots, t_n \rangle \mid \alpha = \text{'tõene'} \}$
- Relatsiooni seisundist korjatakse välja need read, mille puhul α väärtus on tõene.
- α on loogiline avaldis mille argumentideks on relatsiooni R atribuudid ja konstandid.
 - Kasutatakse võrdlustehteid $\{ =, \neq, <, \leq, \geq, > \}$
- Näiteks:

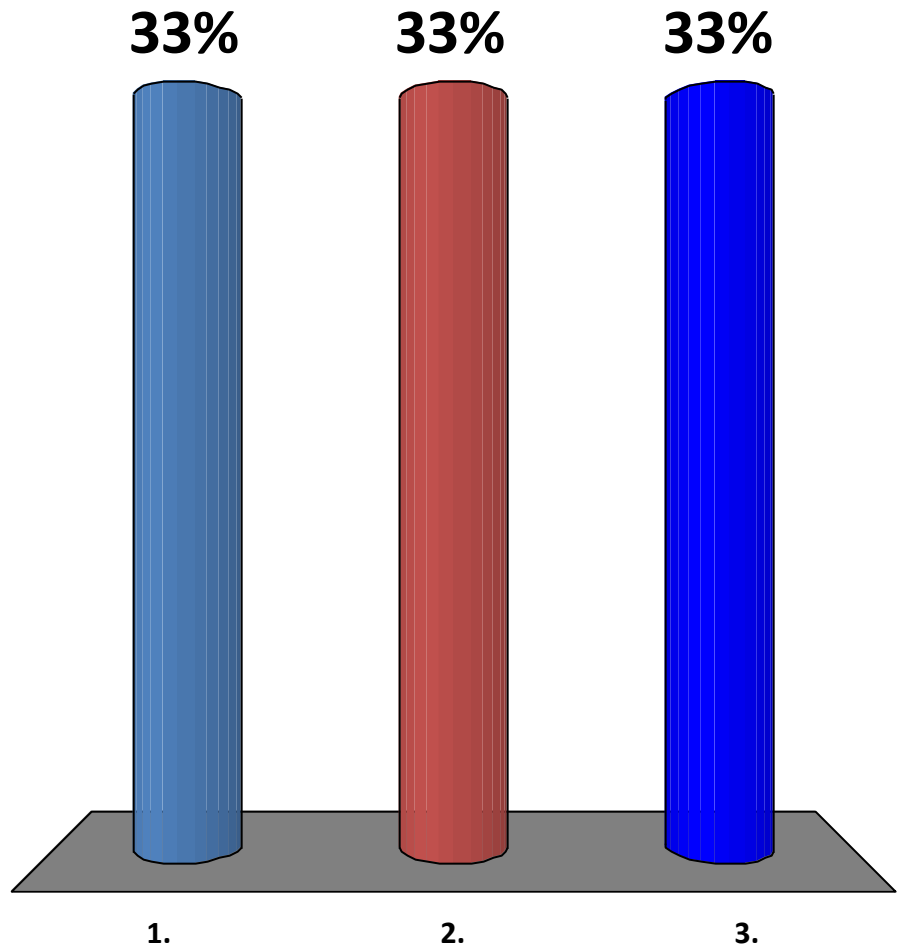
Tudeng (Eesnimi, Pnimi, Matr, aine, Hinne)

$\alpha = (\text{Eesnimi} = \text{'Anne'} \ \& \ \text{Hinne} = \text{'A'})$,

Siis $\sigma_{\alpha} (\text{Tudeng})$ annab tudengite andmed, kelle eesnimi on Anne ja kelle hinne on A.

Olgu R seisundis N rida. Kas filtri operatsiooni tulemuseks saab olla tühi hulk?

1. Jah, saab küll.
2. Ei, ei saa.
3. Ei tea.



Relatsioonialgebra

- Kuna kõigi relatsioonialgebra operatsioonide tulemuseks on relatsioon, saab sellele tulemusele omakorda rakendada teisi relatsioonialgebra operatsioone.
- Analoogia näit. reaalarvude operatsioonidega.
- $\pi_{A_1, A_2}(R) \cup \pi_{B_1, B_3}(S)$
- Tehete järjekord määratud sulgudega
- Unaarsete operatsioonide prioriteet suurem

Näide

- **Kirjanik** (ID, Eesnimi, Perekonnanimi, Riik, Sünnikuupäev)
 - **Väljaanne** (Pealkiri, Aasta, Kirjastus, Lk_arv, Autori_ID, keel)
 - **Riik** (Nimetus, Rahvaarv, GNPpC, Riigikeel)
 - **Kirjastus** (Nimetus, Asukohariik, Aadress)
-
- Kirjanik.ID \equiv Väljaanne.Autori_ID
 - Kirjanik.Riik \equiv Riik.Nimetus
 - **Väljaanne.Kirjastus \equiv Kirjastus.Nimetus**
 - **Kirjastus.Asukohariik \equiv Riik.Nimetus**
 - GNPpC on Gross National Product per Capita
-
- Leida, mis keeles on ingliskeelsete riikide kirjastuste poolt välja antud raamatud.

Lahendus

- $\pi_{\text{keel}} (\sigma_{\alpha} (\text{Riik} \times \text{Kirjastus} \times \text{Väljaanne}))$

$\alpha = \text{Riik.nimetus} = \text{Kirjastus.Asukohariik} \&$

$\text{Kirjastus.Nimetus} = \text{Väljaanne.Kirjastus} \&$

$\text{Riigikeel} = \text{„inglise“}$

Miks on meil vaja relatsioonialgebrat?

- Relatsioonialgebra keele õppimiseks kulub 1-3 tundi.
- On klassikaline näide väga kõrge taseme keelest.
- On andmebaaside keele võimsuse etaloniks.

- NB! Selles keeles ei ole aritmeetikat!
 - S.t. ei saa leida keskmisi, summasid, min/max väärtusi jms.

Kodune ülesanne

- Täisteksti loengumaterjalidest lugeda läbi tuletatud operatsioonid:
 - \oplus -ühend
 - Loomulik ühend
 - Jagamine
- Koostada mõtestatud ülesanne kahele neist.
- Näit. Olgu Laul (Nimi, helistik, viisi-autori_ID)
Helilooja (eesnimi, perenimi, ID, sünd)
Leida kes millise laulu on kirjutanud.