



Eesti Maaülikool
Estonian University of Life Sciences

www.emu.ee

Peakomponentanalüüs (*principal component analysis, PCA*)

Tanel Kaart
Mitmemõõtmelise statistika koolitus
Jaanuar 2016, Tartu

Peakomponentanalüüs (*principal component analysis, PCA*)

Peakomponentanalüüsi eesmärgiks on

- andmemahu vähendamine info koondamise teel,
- latentsete (otseselt mittemõõdetavate) näitajate/mustrite hindamine,
- üldiste indeksite tekitamine,
- objektide/gruppide klasterdamine,
- ...

Tunnused

- Peakomponentanalüüs on rakendatav arvutunnuste korral.
- Tunnused võivad olla
 - pidevad - mõõtmistulemused,
 - diskreetsed - kui mitu midagi on,
 - järjestustunnused - hinnangud näiteks 5-pallisel skaalal,
 - binaarsed (0/1-tüüpi) - midagi toimus/ei toimunud (võivad olla ka nn '*dummy variables*').
- Tunnused ei pea olema mõõdetud samades ühikutes ja/või samal skaalal.
- Analüüsi võib samaaegselt kaasata erinevat tüüpi arvtunnuseid (selles osas võivad erinevate inimeste arvamused lahknedagi ...)

Kui palju uurimisobjekte (n) ja kui palju tunnuseid (k)?

Ühtset vastust pole ☹️😊

- Mõnikord soovitus: $n \gg k$ (a'la 10 mõõtmist iga tunnuse kohta - aga kehtib see soovitus eelkõige faktoranalüüsi puhul).
- Mõnikord väited: 100-200 mõõtmist on igal juhul piisavalt.
- Reaalselt sõltub palju uuringusse kaasatud tunnuste omavaheliste seoste tugevusest, mustrite universaalsusest jmt.
- Võimalik viia läbi nõ ristvalideerimine: teostada peakomponentanalüüs korduvalt, jättes iga kord välja kas üks või mitu vaatlust - kui tulemused erinevate vaatluste välja jätmisel eriti ei muutu, võib lugeda tuvastatud mustrid-indeksid-seaduspärad kehtivateks.

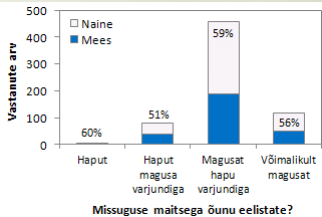
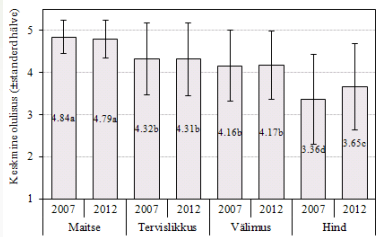
Näide: Ulvi Moor jt, „õunauuring“.

- Ankeetküsitlus aastatel 2007 ja 2012;
- kokku 668 inimest, mehed-naised, viis vanuseklassi, neli haridustaset;
- küsimused kujul
 - „Kas eelistate kodumaist või importõuna“,
 - „Kui olulised on õuna puhul Teile kui tarbijale ...“ jne;
- küsimuste vastused 5-palliskaalal (1-5)

Aasta	Sugu	Vanus	Haridus	Kodumais	Hind	Välimus	Maitse	Tervislikk	Maheus	Magusus
2007	Naine	36-45	Kõrgharidus	5	2	4	5	5	5	3
2007	Naine	36-45	Kõrgharidus	4	4	4	5	5	5	4
2007	Naine	31-35	Keskharidus	4	3	4	4	2	4	2
2007	Naine	36-45	Kõrgharidus	4	5	4	5	5	4	2
2007	Naine	31-35	Kõrgharidus	4	4	4	5	5	2	3
2007	Naine	26-30	Kõrgharidus	4	3	4	5	5	4	4
2007	Naine	36-45	Kõrgharidus	5	4	4	5	5	4	3
2007	Naine	36-45	Kõrgharidus	4	2	4	5	2	2	2

„Õunauuring“ - kuidas uurida seoseid ja erinevusi?

- Tavaline lahendus:
 - sagedustabelid ja vastavad joonised
 - keskmised



- korrelatsioonid

	Kodumaisus	Hind	Välimus	Maitse	Tervislikkus	Maheus	Magusus
Kodumaisus	1						
Hind	0.0008808	1					
Välimus	-0.0786273	0.1235804	1				
Maitse	0.0185897	-0.0942401	0.2282795	1			
Tervislikkus	0.3668458	0.1049378	0.0228645	0.1102184	1		
Maheus	0.3176403	-0.0465162	0.0144182	0.0326808	0.2969469	1	
Magusus	-0.0318697	0.0810915	0.1678925	0.1122004	-0.0066489	-0.0069751	1

„Õunauuring“ - kuidas uurida seoseid ja erinevusi?

- Veel tabeleid ja jooniseid

The first chart shows average scores (standard deviation) for taste and health. The second chart shows average scores for various attributes. The third chart shows the percentage of respondents for different categories.

Kategooria	Väärtus
Mees Maitse	4.76
Naine Maitse	4.86
Mees Tervislikkus	4.12
Naine Tervislikkus	4.46
Mees Välimus	4.10
Naine Välimus	4.50

Kategooria	Väärtus
Kindlasti kodumaist	4.50
Pigem kodumaist	4.83
Päritolu pole oluline	4.79
Pigem importõuna	4.83
Kindlasti importõuna	4.13
Kindlasti kodumaist	4.18
Pigem kodumaist	4.17
Päritolu pole oluline	4.39
Pigem importõuna	4.13
Kindlasti importõuna	4.20
Kindlasti kodumaist	4.34
Pigem kodumaist	4.12
Päritolu pole oluline	3.69
Pigem importõuna	3.44
Kindlasti importõuna	3.53
Keskmine	3.51

Kas eelistate kodumaist või importõuna?

- Kindlasti kodumaist
- Pigem kodumaist
- Päritolu pole oluline
- Pigem importõuna
- Kindlasti importõuna

Kas eelistate mahetootmisest või tavatootmisest pärinevat õuna?

Kategooria	Protsent
Kindlasti mahetootmisest	~65%
Pigem mahetootmisest	~25%
Ei tea, raske öelda	~10%
Pigem tavatootmisest	~15%
Kindlasti tavatootmisest	~5%

Problem:

erinevate tabelite ja jooniste arv on väga suur, mistõttu on raske aduda terviklikku kompleksset pilti ja olulisemaid seaduspärasid!

„Õunauuring“ - peakomponentanalüüs (PCA)

- Eesmärgiks leida algsete tunnuste lineaarkombinatsioonid, mis kirjeldaksid võimalikult hästi ära algseis tunnustes sisalduva info:

The diagram shows the relationship between principal components (PC1, PC2, ..., PC7) and the original variables (x1 to x7). PC1 is associated with x1, x2, x3, x4, x5, x6, and x7. PC2 is associated with x3, x4, x5, and x6. PC7 is associated with x7.

faktoriteks (principal components, factors):

$$PC_1 = b_{11}x_1 + b_{21}x_2 + \dots + b_{71}x_7$$
$$PC_2 = b_{12}x_1 + b_{22}x_2 + \dots + b_{72}x_7$$
$$\dots$$
$$PC_7 = b_{17}x_1 + b_{27}x_2 + \dots + b_{77}x_7$$

- Peakomponentanalüüs moodustab lineaarkombinatsioonid nii, et
 - esimene peakomponent kirjeldab ära võimalikult suure osa kõigi algsete tunnuste koguväruvarieeruvusest,
 - teine võimalikult suure osa alles jäänud varieeruvusest, jne.

„Õunauuring“ - peakomponentanalüüs

- Matemaatiliselt baseerub peakomponentanalüüs uuritavate tunnuste korrelatsiooni- või kovariatsioonimaatriksi **A** omaväärtusanalüüsil:
 - lahendatakse maatriksvõrrand $Ab = \lambda b$ ehk $(A - \lambda I)b = 0$, kus
 - λ on maatriksi **A** omaväärtus (*eigenvalue*), mis peakomponentanalüüsi kontekstis näitab faktori poolt kirjeldatavat varieeruvust,
 - vektor **b** on korrelatsioonimaatriksi omavektor (*eigenvector*),
 - mis peakomponentanalüüsi kontekstis sisaldab korrelatsioonikordajaid algsete tunnuste ja moodustatud peakomponendi vahel ja
 - mille elemendid on ka algsete tunnuste (standardiseeritud) väärtuste kordajaks peakomponendi väärtust määravas võrrandis - nn **faktorlaadungid** (*factor loadings*);
 - kõik omavektorid on omavahel ortogonaalsed, millest tuleneb ka, et vastavad peakomponendid on sõltumatud (nende omavaheline korrelatsioon on null).

„Õunauuring“ - peakomponentanalüüs

Andmed

Aasta	Sugu	Vanus	Häridus	KodumaisHind	Valimus	Maitse	Tervislikk	Maheus	Magusus
2007	Naine	36-45	Kõrgharidus	5	2	4	5	5	3
2007	Naine	36-45	Kõrgharidus	4	4	4	5	5	4
2007	Naine	31-35	Keskharidus	4	3	4	4	2	4
2007	Naine	36-45	Kõrgharidus	4	5	4	5	5	4
2007	Naine	31-35	Kõrgharidus	4	4	4	5	5	2
2007	Naine	26-30	Kõrgharidus	4	3	4	5	5	4
2007	Naine	36-45	Kõrgharidus	5	4	4	5	5	4

Korrelatsiooni-maatriks

	Kodumaisus	Hind	Valimus	Maitse	Tervislikkus	Maheus	Magusus
Kodumaisus	1.0000	0.0024	-0.1168	0.0390	0.3672	0.3150	-0.0336
Hind	0.0024	1.0000	0.0969	-0.0955	0.0323	-0.0532	0.0803
Valimus	-0.1168	0.0969	1.0000	0.2488	-0.0026	0.0009	0.1644
Maitse	0.0390	-0.0955	0.2488	1.0000	0.1175	0.0463	0.1065
Tervislikkus	0.3672	0.0323	-0.0026	0.1175	1.0000	0.3010	-0.0253
Maheus	0.3150	-0.0532	0.0009	0.0463	0.3010	1.0000	-0.1037
Magusus	-0.0336	0.0803	0.1644	0.1065	-0.0253	-0.1037	1.0000

Omaväärtused

	Eigenvalue	Difference	Proportion	Cumulative
1	1.68647373	0.31979660	0.2409	0.2409
2	1.36667713	0.27516894	0.1952	0.4362
3	1.09150819	0.21761178	0.1559	0.5921
4	0.87389641	0.11057473	0.1248	0.7169
5	0.76332168	0.13463642	0.1090	0.8260
6	0.62862526	0.03912765	0.0898	0.9158
7	0.58349761		0.0842	1.0000

Omavektorid

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Kodumaisus	0.586188	-0.102402	0.049232	0.130467	-0.232729	-0.646768	0.392403
Hind	0.003733	0.150112	0.853858	-0.279199	-0.090420	0.223958	0.334805
Valimus	-0.064445	0.651881	0.012185	-0.320255	0.466492	-0.479779	-0.142776
Maitse	0.138294	0.550956	-0.441186	-0.227814	-0.451197	0.329749	0.344188
Tervislikkus	0.582203	0.080325	0.150378	-0.173669	-0.237079	0.136212	-0.725984
Maheus	0.539543	0.009267	-0.085812	0.152815	0.669466	0.415805	0.238774
Magusus	-0.054989	0.491503	0.209123	0.034896	-0.109897	0.034131	-0.104925

„Õunauuring“ - peakomponentanalüüs

Omapäärtused

	Eigenvalue	Difference	Proportion	Cumulative
1	1.68647373	0.31979660	0.2409	0.2409
2	1.36667713	0.27516894	0.1952	0.4362
3	1.09150819	0.21761178	0.1559	0.5921
4	0.87389641	0.11057473	0.1248	0.7169
5	0.76332168	0.13469642	0.1090	0.8260
6	0.62862526	0.03912765	0.0898	0.9158
7	0.58949761		0.0842	1.0000

- Korrelatsioonimaatriksi **omaväärtused** näitavad peakomponendi poolt kirjeldatavat varieeruvust ja selle osa koguvarieeruvusest.
- Et algsed tunnused analüüsi eel **standardiseeritakse**, on nende dispersioon võrdne ühega ja sestap omavad üksiktunnustest suuremat kirjeldusvõimet vaid peakomponendid, millele vastav omaväärtus > 1.

$$st(x_1) = \frac{x_1 - \bar{x}_1}{S_{x_1}}$$

„Õunauuring“ - peakomponentanalüüs

Omapäärtused

- Erinevad programmid väljastavad tulemusi erineval kujul!

	Eigenvalue	Difference	Proportion	Cumulative
1	1.68647373	0.31979660	0.2409	0.2409
2	1.36667713	0.27516894	0.1952	0.4362
3	1.09150819	0.21761178	0.1559	0.5921
4	0.87389641	0.11057473	0.1248	0.7169
5	0.76332168	0.13469642	0.1090	0.8260
6	0.62862526	0.03912765	0.0898	0.9158
7	0.58949761		0.0842	1.0000

```
> apple.pc.cor <- princomp(appledata[,c(-1:-4)], cor=TRUE, scores=TRUE)
> summary(apple.pc.cor)
Importance of components:
      Comp.1  Comp.2  Comp.3  Comp.4  Comp.5  Comp.6  Comp.7
Standard deviation  1.2960971 1.1713805 1.0450556 0.9372917 0.8733405 0.78392800 0.7746460
Proportion of Variance 0.2399811 0.1960189 0.1560202 0.1255022 0.1089605 0.08779187 0.0857252
Cumulative Proportion 0.2399811 0.4360000 0.5920202 0.7175224 0.8264829 0.91427480 1.0000000

> apple.pc.cor$sdev^2 # peakomponentide dispersioonid (omaväärtused)
      Comp.1  Comp.2  Comp.3  Comp.4  Comp.5  Comp.6  Comp.7
1.6798677 1.3721324 1.0921411 0.8785157 0.7627236 0.6145431 0.6000764
```

R-i funktsioon *princomp*

```
> apple.pc.cor2 = dudi.pca(appledata[,c(-1:-4)], scale=T, scannf=F)
> summary(apple.pc.cor2)
Class: pca dudi
Call: dudi.pca(df = appledata[, c(-1:-4)], scale = T, scannf = F)

Total inertia: 7

Eigenvalues:
  Ax1  Ax2  Ax3  Ax4  Ax5
1.6799 1.3721 1.0921 0.8785 0.7627

Projected inertia (%):
  Ax1  Ax2  Ax3  Ax4  Ax5
24.00 19.60 15.60 12.55 10.90

Cumulative projected inertia (%):
  Ax1 Ax1:2 Ax1:3 Ax1:4 Ax1:5
24.00 43.60 59.20 71.75 82.65
```

R-i funktsioon *dudi.pca*

SAS-i protseduur *princomp*

„Õunauuring“ - peakomponentanalüüs

Omavektorid

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Kodumaisus	0.586188	-0.102402	0.049232	0.130467	-0.232729	-0.646768	0.392403
Hind	0.003733	0.150112	0.853858	-0.279199	-0.090420	0.223958	0.334805
Välimus	-0.064445	0.651881	0.012185	-0.320255	0.466492	-0.479779	-0.142776
Maitse	0.138294	0.550956	-0.441186	-0.227814	-0.451197	0.329749	0.344188
Tervislikkus	0.582203	0.080325	0.150378	-0.173669	-0.237079	0.136212	-0.725394
Maheusus	0.539543	0.009267	-0.085812	0.152815	0.669466	0.415805	0.238774
Magusus	-0.054988	-0.481603	0.209123	0.834896	-0.109897	0.034131	-0.104925

- Korrelatsioonimaatriksi omavektorid (*eigenvectors*) kujutavad enesest korrelatsioonikordajaid algsete tunnuste ja moodustatud peakomponentide vahel.
- Samuti on omavektori elemendid algsete tunnuste standardiseeritud väärtuste kordajaks peakomponenti väärtust määravas võrrandis - nn faktorlaadungid (*factor loadings*).

Näiteks avaldub esimese peakomponenti väärtus kujul

$$PC_1 = 0,586 \cdot \text{st(Kodum.)} + 0,004 \cdot \text{st(Hind)} - 0,064 \cdot \text{st(Välimus)} + \dots$$

siin st() tähendab standardiseeritud tunnust.

„Õunauuring“ - peakomponentanalüüs

Omavektorid

- Erinevad programmid võivad anda tulemuseks vastupidise märgiga omavektorid!

Näiteks antud juhul $PC2_R = -PC2_{SAS}$.

R:

```
> apple.pc.cor2$c1
```

	CS1	CS2
Kodumaisus	0.586188488	0.102402077
Hind	0.003732834	-0.15011635
Välimus	-0.064444921	-0.651880722
Maitse	0.138293607	-0.550955667
Tervislikkus	0.582203193	-0.080325460
Maheusus	0.539542991	-0.009266879
Magusus	-0.054987774	-0.481603192

SAS:

	PC1	PC2
Kodumaisus	0.586188	-0.102402
Hind	0.003733	0.150112
Välimus	-0.064445	0.651881
Maitse	0.138294	0.550956
Tervislikkus	0.582203	0.080325
Maheusus	0.539543	0.009267
Magusus	-0.054988	0.481603

„Õunauuring“ - peakomponentanalüüs

Skoorid

- Peakomponentide väärtuseid nimetatakse ka **skoorideks (scores)**.

Skoorid on välja arvutatavad iga andmebaasi rea tarvis:

Kodum:	Hind	Valim:	Maitse	Tervisli:	Mahe:	Magusi:	PCA1	PCA2	PCA3	PCA4	PCA5
5	2	4	5	5	5	5	1.901704	0.264187	0.399534	-0.32087	-0.52952
3		4	5	4	4	4					
4		4	5	5	5	5	0.771246	1.260386	0.168522	1.385335	-0.0863
4		4	4	2	4	2	-0.38148	-3.14741	-0.01897	-0.00201	1.475696
4	5	4	5	5	4	2	-0.38591	0.426845	-0.43400	-0.00000	0.00000
4	4	4	5	5	2	3	-0.76525	-0.05342			
4	3	4	5	5	4	4	1.011838	0.40264			

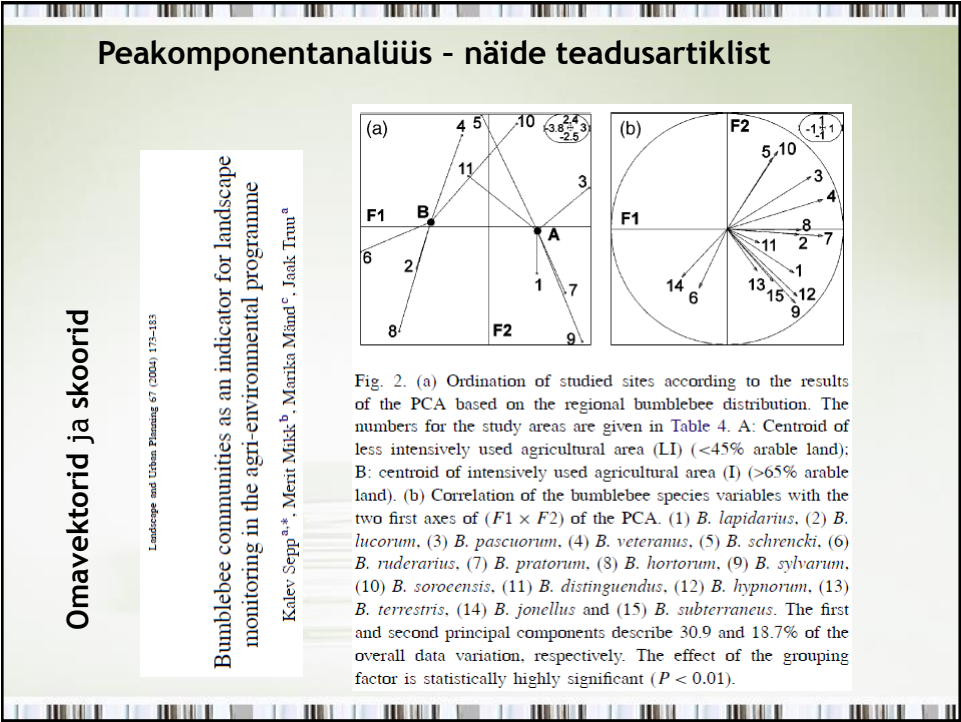
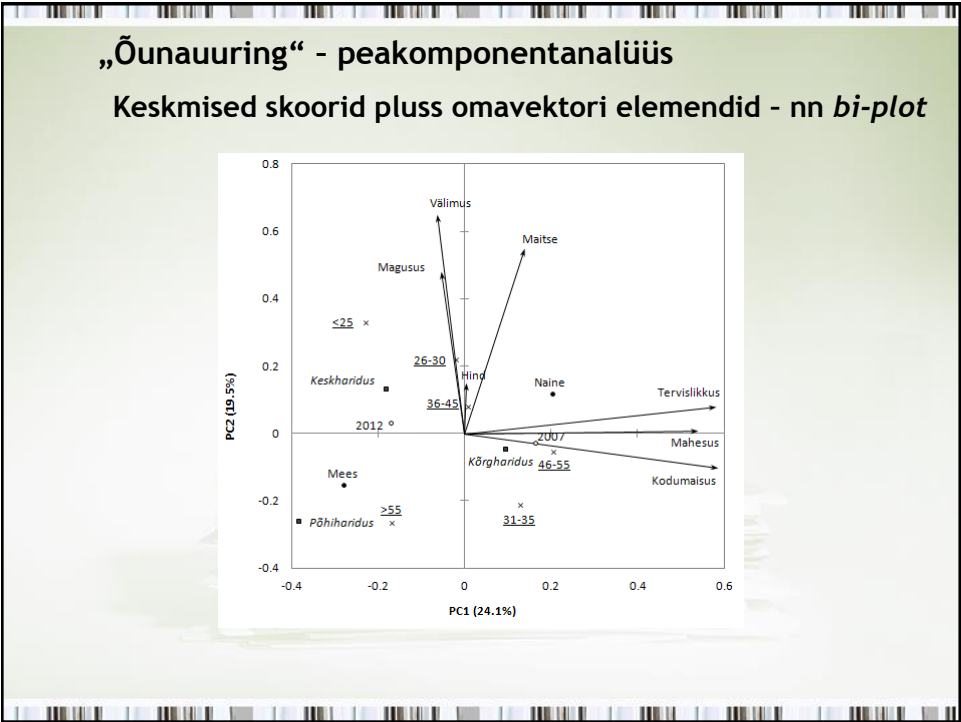
- NB! Kui mõned uurimisobjektid on mõne algse tunnuse väärtus puudu, jääb taoline objekt vaikumisi analüüsist välja (st, et ei arvestata ka tema teisi, olemasolevaid väärtuseid).

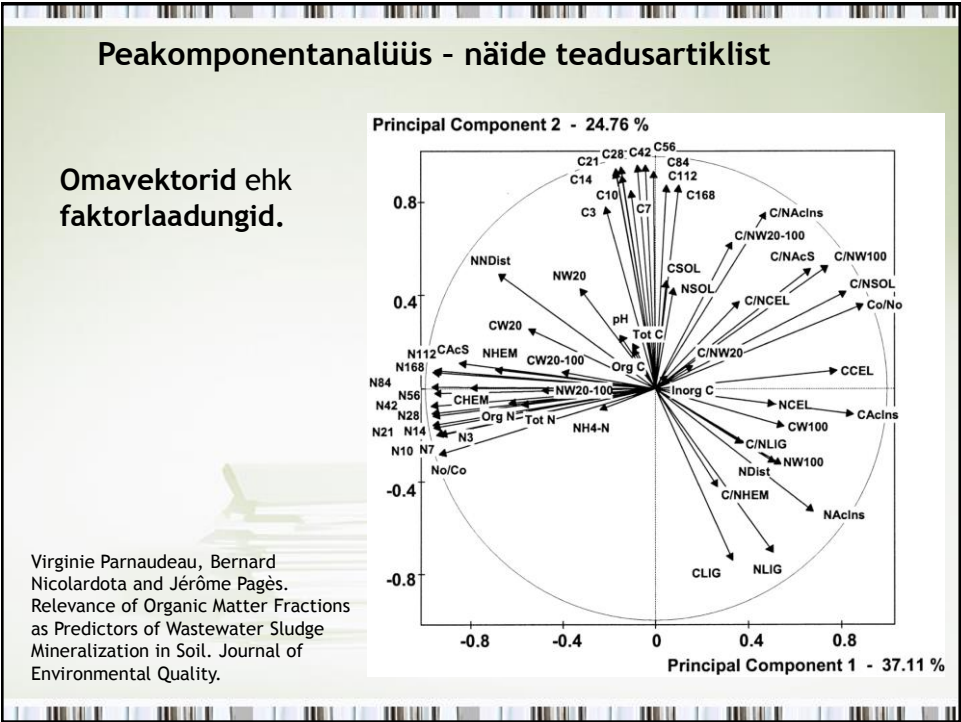
„Õunauuring“ - peakomponentanalüüs

Skoorid

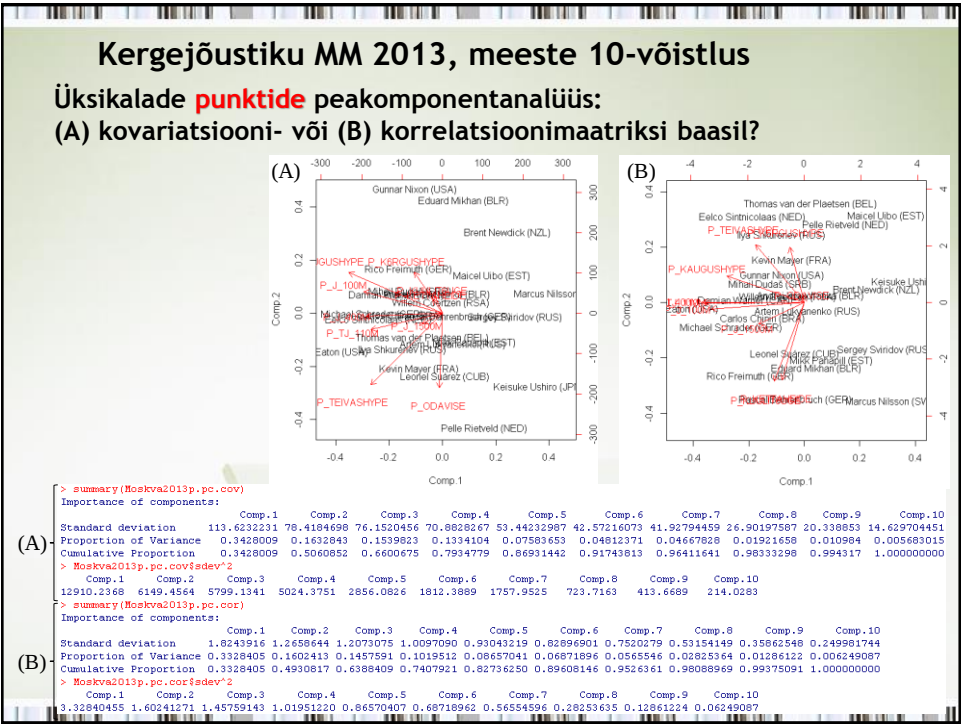
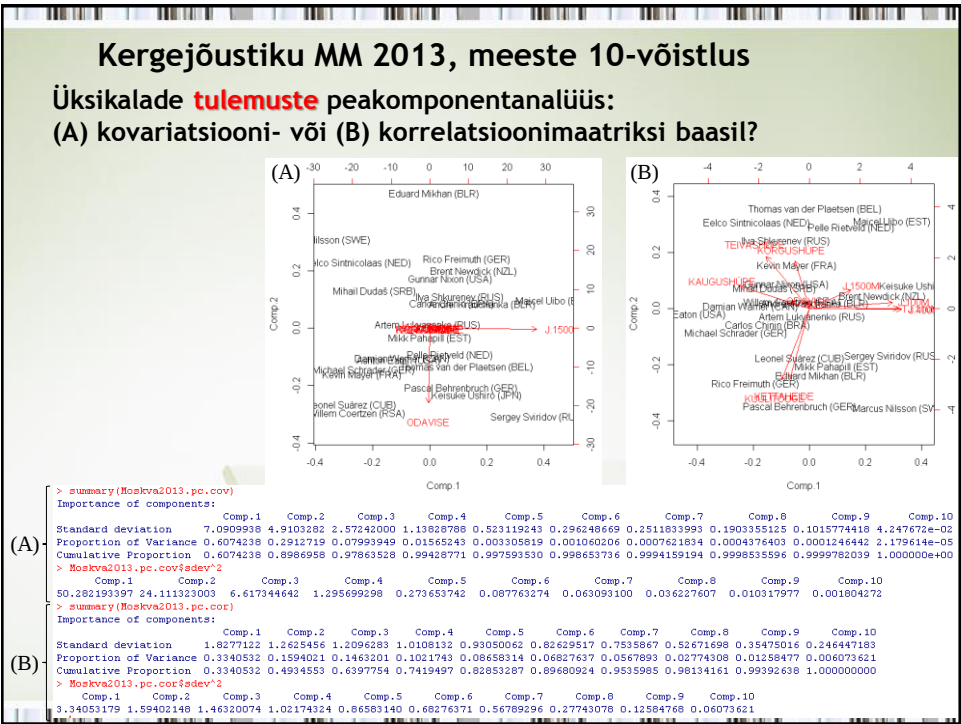
- Peakomponentide väärtused e skoorid võimaldavad peakomponente seostada teiste andmebaasi kuuluvate tunnustega ja teha seeläbi üldisemaid järeldusi andmetes leiduvate mustrite/struktuuride kohta.

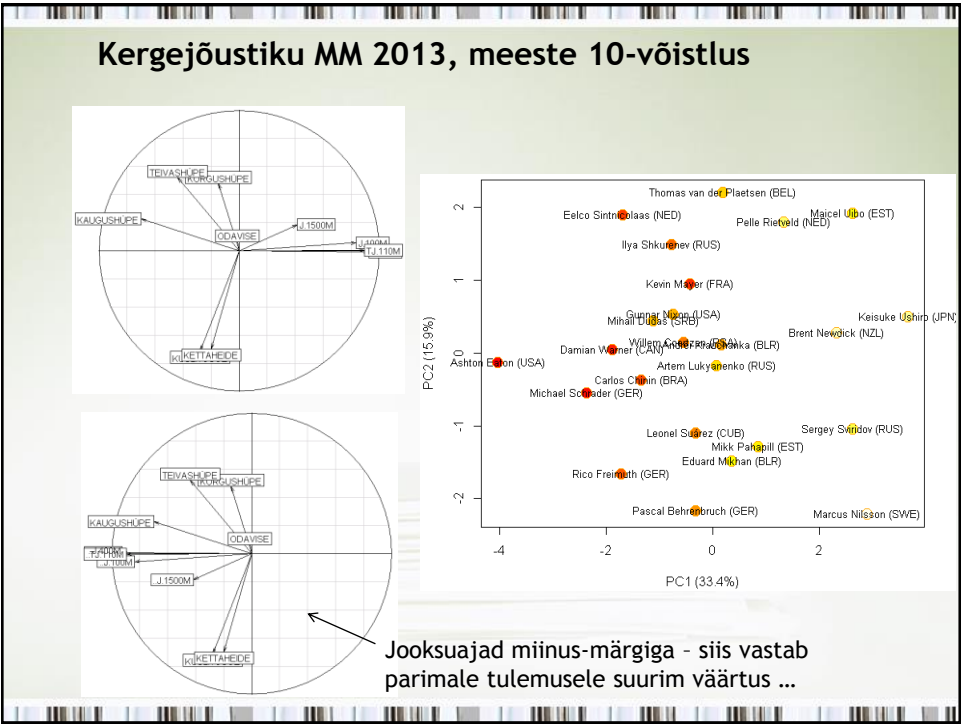
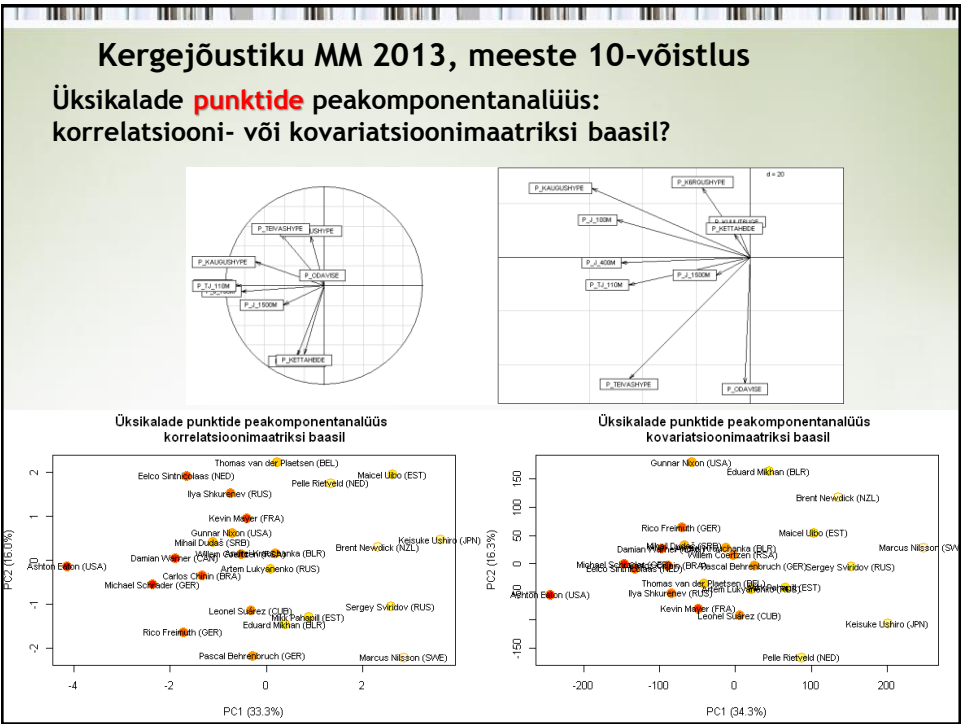
Samas ei oma peakomponentide väärtused mingit sisulist tähendust arvestamata nende seotust algsete tunnustega.



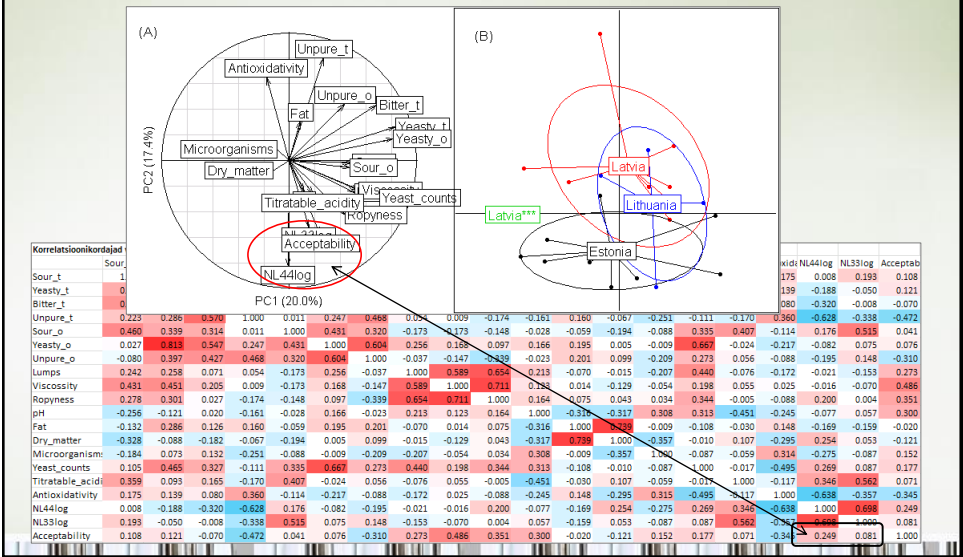


Kergejõustiku MM 2013, meeste 10-võistlus										
NIMI	KOHT	KOKKU	J 100M	KAUGUSH	KUULITÕU	KÕRGUSH	J 400M	TJ 110M	KETTAHEIITEIVASHÜ	ODAVISE J 1500M
Ashton Eaton (USA)	1	8809	10.35	7.73	14.39	1.93	46.02	13.72	45	5.2 64.83 269.8
Michael Schrader (GER)	2	8670	10.73	7.85	14.56	1.99	47.66	14.29	5	65.67 265.4
Damian Warner (CAN)	3	8512	10.43	7.39	14.23	2.05	48.41	13.96	44.13	4.8 64.67 270
Kevin Mayer (FRA)	4	8446	11.23	7.5	13.76	2.05	49.53	14.21	45.37	5.2 66.09 265
Eelco Sintnicolaas (NED)	5	8391	10.85	7.65	14.08	2.02	48.25	14.18	39.21	5.3 56.75 264.6
Carlos Chinin (BRA)	6	8388	10.78	7.54	14.49	1.96	48.8	14.05	45.84	5.1 59.98 276
Rico Freimuth (GER)	7	8382	10.6	7.22	14.8	1.99	48.05	13.9	48.74	4.9 56.21 277.8
Ilya Shkurennev (RUS)	8	8370	10.97	7.35	13.88	2.05	48.39	14.34	44.06	5.4 59.46 277
Willem Coertzen (RSA)	9	8343	10.95	7.44	13.88	2.05	48.32	14.3	43.25	4.5 69.35 264.6
Leonel Suárez (CUB)	10	8317	11.07	7.33	14.2	1.9	48.21	14.62	46.41	4.9 68.61 263.9
Pascal Behrenbruch (GER)	11	8316	10.95	7.19	15.86	1.99	48.4	14.46	45.66	4.7 67.07 277.2
Andrei Krauchanka (BLR)	12	8314	11.19	7.39	14.84	2.11	49.65	14.44	46.12	5.1 59.98 279.6
Gunnar Nixon (USA)	13	8312	10.84	7.8	14.68	2.14	48.56	14.57	42.38	4.6 57.97 275.8
Mihail Dudaș (SRB)	14	8275	10.67	7.51	13.45	1.96	47.73	14.59	44.06	4.9 59.06 266.6
Thomas van der Plaetsen (BEL)	15	8255	11.09	7.64	13.57	2.05	49	14.66	41.17	5.1 65.31 277.9
Artem Lukyanenko (RUS)	16	8177	11.11	7.09	14.43	1.99	49.01	14.21	44.06	5 61.83 273
Mikk Pahapill (EST)	17	8170	11.33	7.12	14.87	1.99	49.84	14.54	47.57	5 62.89 273.2
Eduard Mikhan (BLR)	18	7968	10.97	7.42	14.15	1.96	48.8	14.82	48	4.6 50.74 273.7
Maicel Uibo (EST)	19	7850	11.21	7.26	13.68	2.11	50.6	15.29	41.84	4.9 59.63 288.3
Sergey Sviridov (RUS)	20	7843	10.98	7.3	14.2	1.9	51.13	15.23	45.26	4.5 69.38 286.4
Pelle Rietveld (NED)	21	7840	11.15	6.68	13.21	1.9	48.67	14.37	38.06	5.1 64.38 275.9
Keisuke Ushiro (JPN)	22	7751	11.44	6.84	13.17	1.96	51.2	15.32	45.31	4.9 67.65 278.9
Brent Newdick (NZL)	23	7744	11.14	7.28	13.84	2.02	50.25	15.19	42.49	4.5 57.3 278.5
Marcus Nilsson (SWE)	24	7540	11.42	6.62	14.81	1.87	50.56	15.04	42.49	4.5 54.86 260.1





Näide: Piret Raudsepp jt, „keefiride uuring“.
Peakomponentanalüüsi tulemus ei pruugi ühtida
korrelatsioonanalüüsi tulemustega!

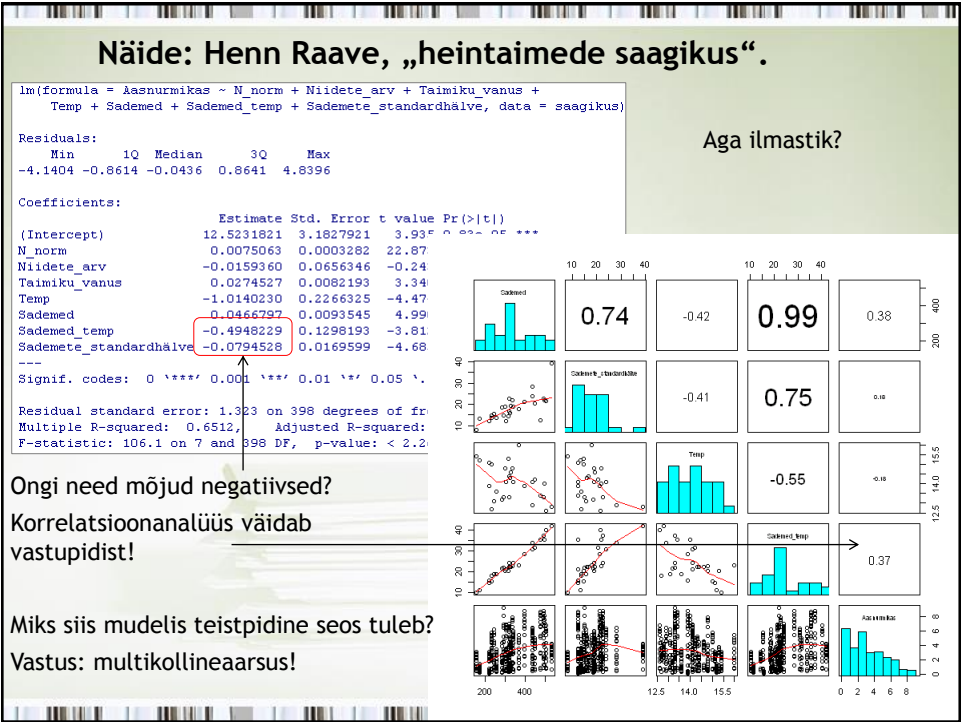


Näide: Henn Raave, „heintaimede saagikus“.
Peakomponentide väärtuste (e skooride) kasutamine
modelleerimisel (näiteks *principal component regression*)

N_norm	Niide	a	Katse	aast	Taimiku	v	Sademed	Sademet	Temp	Sademed	Kerahein	Aasnurmil	Teised	Kogusaak
0	5	1971	1	219.5	12.2	13.9	15.8	0.97	0.04	0.08	1.08			
0	5	1972	2	316.5	14.5	15.3	20.6	0.23	0.38	0.58	1.18			
0	5	1973	3	340.1	18.9	14.1	24.1	0.38	0.32	0.78	1.48			
0	5	1974	4	291.4	14.8	13.3	22	0.43	0.71	0.02	1.17			
0	5	1975	5	219.2	12.3	15.1	14.5	0.44	0.68	0	1.12			
0	5	1976	6	231.8	13.5	12.7	18.3	0.24	0.3	0.13	0.67			
0	5	1977	7	284.9	18.5	13.2	21.5	0.16	0.12	0.14	0.42			
0	5	1978	8	522.5	29.5	17.8	41.8	0.46	0.6	0.06	1.12			

```
lm(formula = Aasnurmikas ~ N_norm + Niide
```

Aga ilmastik?



Näide: Henn Raave, „heintaimede saagikus“.

Mida teha?

- Mitte pöörata tähelepanu ilmastikunäitajate mõju olemusele - eesmärgiks ei ole hinnata ilmastiku mõju, vaid elimineerida ilmastiku mõju tegelikult huvi pakkuvate faktorite mõjust ja/või hinnata, kuivõrd kogu faktorite komplekt saagikust üldiselt prognoosida võimaldab.
- Valida välja vaid üks ilmastikunäitaja ja konstrueerida mudel sellega või konstrueerida mitu alternatiivset mudelit, igas erinev ilmastikunäitaja.
- Rakendada omavahel seoses olevatele ilmastikunäitajatele peakomponentanalüüsi ning kasutada saagikust prognoosivas mudelis esimest (ja vajadusel ka teist-kolmandat) peakomponenti; järeldused ilmastikunäitajate mõjust saagikusele tulenevad siis peakomponenti seosest algsete tunnustega „vaadatuna läbi“ peakomponentidele vastava regressioonikordaja mudelis.

