

Interpreting and storing collected data

Katalin T. Biró

[tbk@ace.hu](mailto:tbk@ace.hu)

K. T. Biró, EPISCON course 24/01/2007

What do we already know about...

Data analysis?

Archaeological data?

Analytical data?

K. T. Biró, EPISCON course 24/01/2007

## Data

numbers	nominal variables (name, tel. etc.)
words	ordinal variables (rank)
images	interval scale variables ( $^{\circ}$ C)
	numerical variables (+,-,x,:)

K. T. Biró, EPISCON course 24/01/2007

## Database

Structured data  
model: text-based  
hierarchical  
relational (RDBMs)

K. T. Biró, EPISCON course 24/01/2007

## Database

text-based - string search, Boolean operators

and

or

>

<

typical example: [google](#)

no structure needed

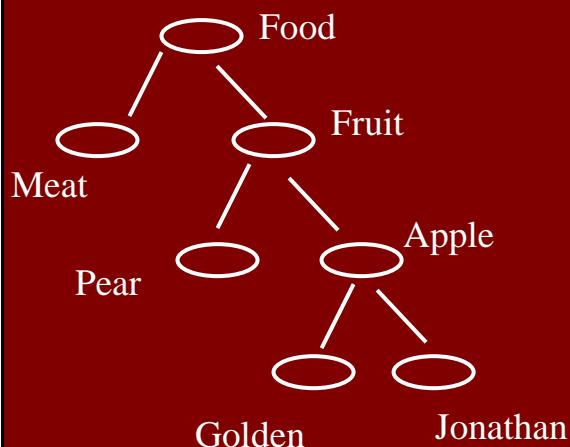
lot of irrelevant data

sensitive to mistakes

K. T. Biró, EPISCON course 24/01/2007

## Database

hierarchical - networking model



Human brain works  
this way

excellent for  
personal databases

not applicable for  
shared information  
systems

K. T. Biró, EPISCON course 24/01/2007

## Database

relational (RDBMs) - data tables connected

simple tables

relations

1 to 1

1 to many

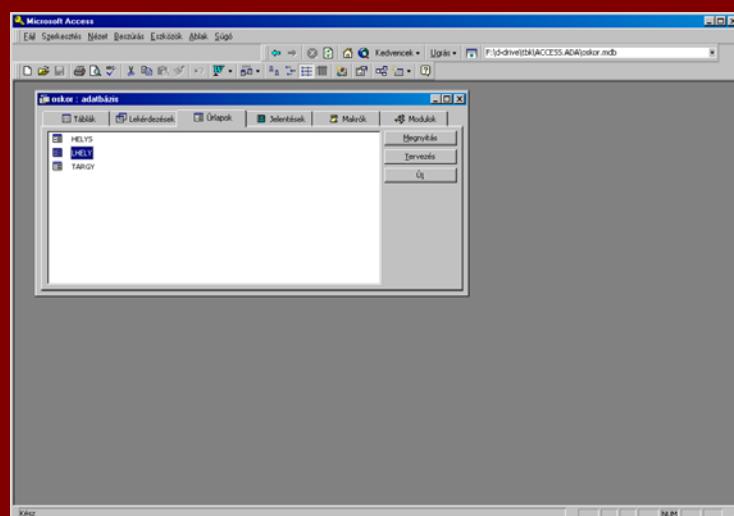
many to many

### example

K. T. Biró, EPISCON course 24/01/2007

## Database

relational (RDBMs) - data tables connected



K. T. BIRÓ, EPISCON course 24/01/2007

## Database

relational (RDBMs) - data tables connected

The screenshot shows a Microsoft Access database window with the 'LHELY' table open. The table has 12 columns: Azonosító, Id, Lelhely neve, Lelhely neve Pontos megne, Ország, MEGY\_KOD, Megye, Műpoint, and gyűjtőkód. The data includes various Hungarian locations like Budapest, Székesfehérvár, Szombathely, etc., across different counties and districts.

Azonosító	Id	Lelhely neve	Lelhely neve Pontos megne	Ország	MEGY_KOD	Megye	Műpoint	gyűjtőkód
1	1	361 Budapest	Angyalföld	HUN	1	Budapest	0	0
2	2	1663 Mezőkövesd	Klementina	HUN	6	Borsod-Abaúj-Z	0	+
3	3	342 Arad	Zomán puszta	HUN	9	Hajdú-Bihar	0	+
4	4	248 Székesfehérvár	Szelidő komp	HUN	5	Békés	0	+
5	5	314 Bodrogkeresztúr	Ung	HUN	6	Borsod-Abaúj-Z	0	+
6	6	314 Bodrogkeresztúr	Székiány	HUN	5	Borsod-Abaúj-Z	0	+
7	7	2644 Tát	Temető mellel	HUN	11	Komárom-Esztergom	0	+
8	8	2637 Tárnok		HUN	13	Pest	0	+
9	9	1743 Nagybékés	Szalacske-2. ár	HUN	14	Somogy	0	+
10	10	2743 Székelyváros	Ung	HUN	10	Heves	0	+
11	11	623 Dunajászváros	Duna	HUN	7	Fehér	0	+
12	12	2462 Széchelmélt	Téglaúgyár	HUN	13	Pest	0	+
13	13	2646 Tatabánya	Mészáros u.	HUN	11	Komárom-Esztergom	0	+
14	14	2477 Szeghalom	Tóviskes Alt. G.	HUN	4	Békés	0	+
15	15	610 Dunaharaszti		HUN	13	Pest	0	+
16	16	2527 Gyöngyös		HUN	16	Csongrád	0	+
17	17	1527 Mát	Garambóz-608	HUN	6	Borsod-Abaúj-Z	0	+
18	18	2623 Tarcal	Mandulás-dűbő	HUN	5	Borsod-Abaúj-Z	0	+
19	19	2960 Viss	Kuttas	HUN	5	Borsod-Abaúj-Z	0	+
20	20	816 Gávavencsellő	Zug	HUN	15	Szabolcs-Szat	0	+
21	21	2715 Tiszalúc	Dankadomb	HUN	5	Borsod-Abaúj-Z	0	+
22	22	2122 Prugy	Tökföl	HUN	5	Borsod-Abaúj-Z	0	+
23	23	3324 Székelyváros	Ung	HUN	0	+	0	+
24	24	3070 Csíkszereda	1. felület 1. gld	HUN	20	Zala	0	+
25	25	1666 Mezőszombor		HUN	5	Borsod-Abaúj-Z	0	+
26	26	2196 Rakamaz	Timár közötti	HUN	15	Szabolcs-Szat	0	+
27	27	2900 Nyírbogd		HUN	15	Szabolcs-Szat	0	+
28	28	301 Gyomarnémeti	Endrőd határát	HUN	4	Békés	0	+
29	29	319 Felgyő	Vitrapart	HUN	6	Csongrád	0	+
30	30	719 Felgyő		HUN	0	+	0	+

K. T. Biró, EPISCON course 24/01/2007

## Database

relational (RDBMs) - data tables connected

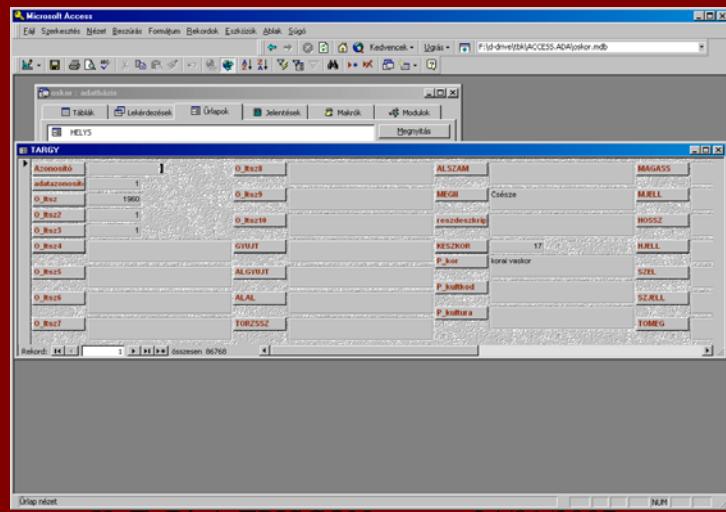
The screenshot shows a Microsoft Access database window with the 'TARGY' table open. The table has 9 columns: TJEELL, P\_Ihked, GYORSZ, GYHELY, P\_Ilehetlen, GYPOINTOS, P\_reteg, GY\_MOD, and GM\_NEV. The data includes various subjects taught in different schools across Hungary.

TJEELL	P_Ihked	GYORSZ	GYHELY	P_Ilehetlen	GYPOINTOS	P_reteg	GY_MOD	GM_NEV	R_elhelyezes
1	HUN	361	Budapest	Angrallid					
1	HUN	361	Budapest	Angrallid					
1	HUN	361	Budapest	Angrallid					
1	HUN	361	Budapest	Angrallid					
1	HUN	361	Budapest	Angrallid					
1	HUN	361	Budapest	Angrallid					
1	HUN	361	Budapest	Angrallid					
1	HUN	361	Budapest	Angrallid					
1	HUN	1663	Mezőkövesd	Klementina					
2	HUN	1663	Mezőkövesd	Klementina					
2	HUN	1663	Mezőkövesd	Klementina					
2	HUN	1663	Mezőkövesd	Klementina					
2	HUN	1663	Mezőkövesd	Klementina					
2	HUN	1663	Mezőkövesd	Klementina					
3	HUN	94	Arad	Zomán puszta					
3	HUN	94	Arad	Zomán puszta					
3	HUN	94	Arad	Zomán puszta					
3	HUN	94	Arad	Zomán puszta					
3	HUN	94	Arad	Zomán puszta					
3	HUN	94	Arad	Zomán puszta					
3	HUN	94	Arad	Zomán puszta					
3	HUN	94	Arad	Zomán puszta					
3	HUN	94	Arad	Zomán puszta					
3	HUN	94	Arad	Zomán puszta					
3	HUN	94	Arad	Zomán puszta					

K. T. Biró, EPISCON course 24/01/2007

## Database

relational (RDBMs) - data tables connected

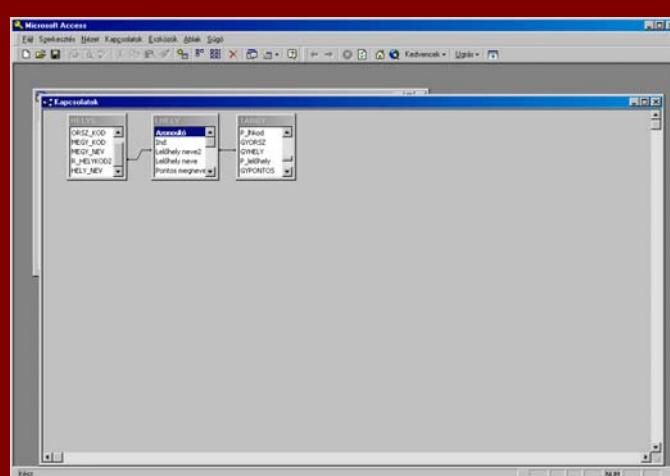


A screenshot of Microsoft Access showing a linked table view. The main window displays a grid of data from two tables: 'HELYS' and 'TARGY'. The 'HELYS' table has columns 'Adressz', 'adatazonosito', 'O\_Nev', 'ALSCZAM', and 'MAGASS'. The 'TARGY' table has columns 'GYUJT', 'RESZTEROK', 'P\_kor', 'korai\_vektor', 'AL\_GYUJT', 'P\_Jellem', 'P\_Jeltra', and 'TOMEQ'. A relationship line connects the 'O\_Nev' column of the 'HELYS' table to the 'GYUJT' column of the 'TARGY' table. The status bar at the bottom shows 'Record: 14 of 14'.

K. T. Biró, EPISCON course 24/01/2007

## Database

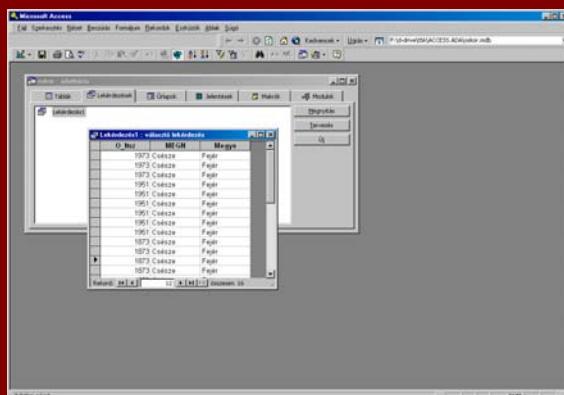
relational (RDBMs) - data tables connected



K. T. Biró, EPISCON course 24/01/2007

Database

relational (RDBMs) - data tables connected



K. T. Biró, EPISCON course 24/01/2007

## Query

## SQL: structured query language

## Boolean operators

K. T. Biró, EPISCON course 24/01/2007

## Statistical evaluation

### Basic concepts

mean  
median  
mode  
variance  
standard deviation

K. T. Biró, EPISCON course 24/01/2007

## Statistical evaluation

### References

THOMAS 1986 Thomas, David Hurst Refiguring Anthropology / First principles of probability & statistics Prospect Heights, Ill. USA Waveland Press Inc. 1986, 1-532

Exploratory Data Analysis. Engineering statistics handbook (URL)

K. T. Biró, EPISCON course 24/01/2007

## Statistical evaluation

Tools:

- Excel: the pedestrian way to almost everything
  - SPSS
  - JMP
  - SAS
- many more

K. T. Biró, EPISCON course 24/01/2007

## Spatial distribution

Surfer

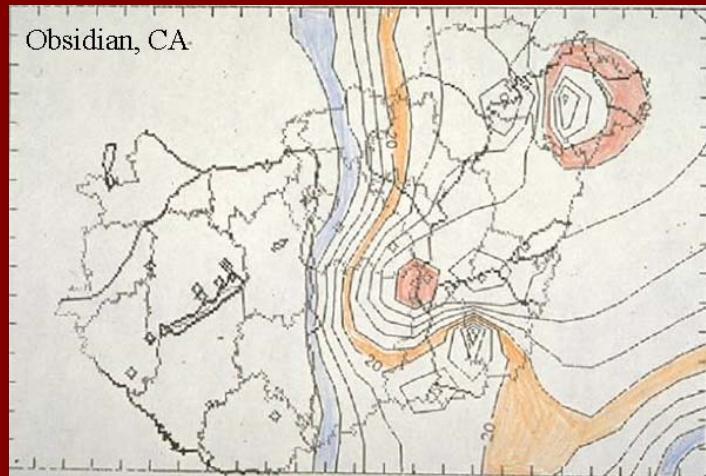
GIS softwares

MAPINFO

K. T. Biró, EPISCON course 24/01/2007

## Spatial distribution

Archaeological distribution of obsidian in Hungary



K. T. Biró, EPISCON course 24/01/2007

## Spatial distribution

Archaeological distribution of obsidian in Hungary

Raw data:

- site name
- coord-lat
- coord-long
- chronological phase
- obsidian pieces
- obsidian %
- stones total

K. T. Biró, EPISCON course 24/01/2007

## Simple techniques

Histograms

Stem-and-leaf plot

Box and whiskers plot

Crossplots

K. T. Biró, EPISCON course 24/01/2007

## Multivariate techniques

Principal component analysis

Cluster analysis

K. T. Biró, EPISCON course 24/01/2007

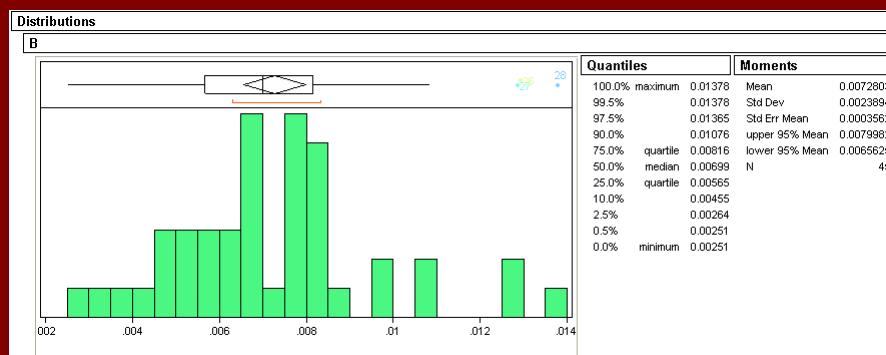
## Case study: PGAA of grey flint

### original data table

K. T. Biró, EPISCON course 24/01/2007

## Case study: PGAA of grey flint

### histogram with box-and-whiskers



### all histograms

K. T. Biró, EPISCON course 24/01/2007

## Case study: PGAA of grey flint stem-and-leaf plot

SiO<sub>2</sub>

98,430061

98,1616863

98,3779764

98,3933189

98,3621029

98,3152068

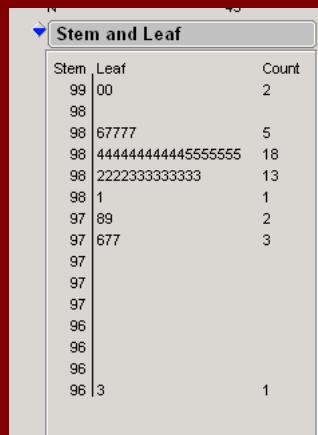
98,4703701

98,3239508

98,6514208

98,450583

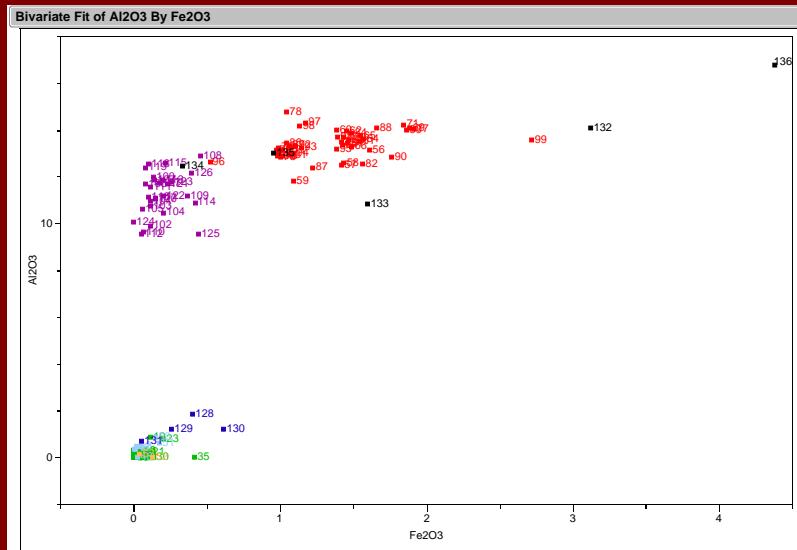
98,5340224



K. T. Biró, EPISCON course 24/01/2007

## Case study: PGAA of grey flint crossplot

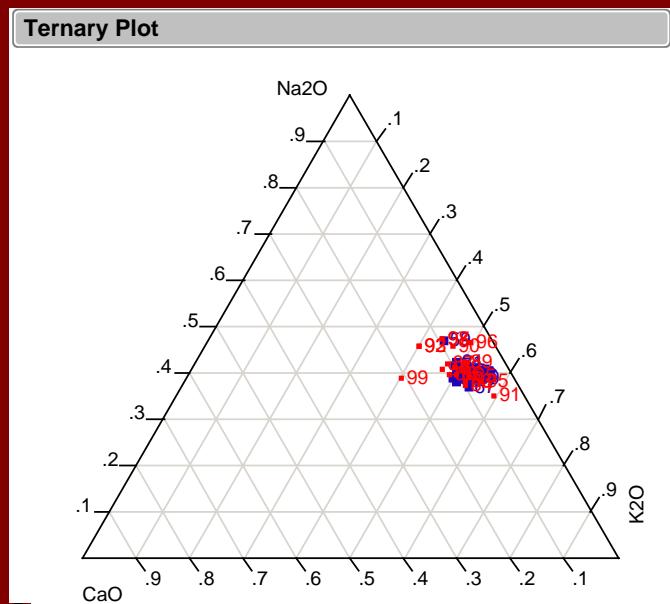
---



K. T. BIRO, EPISCON course 24/01/2007

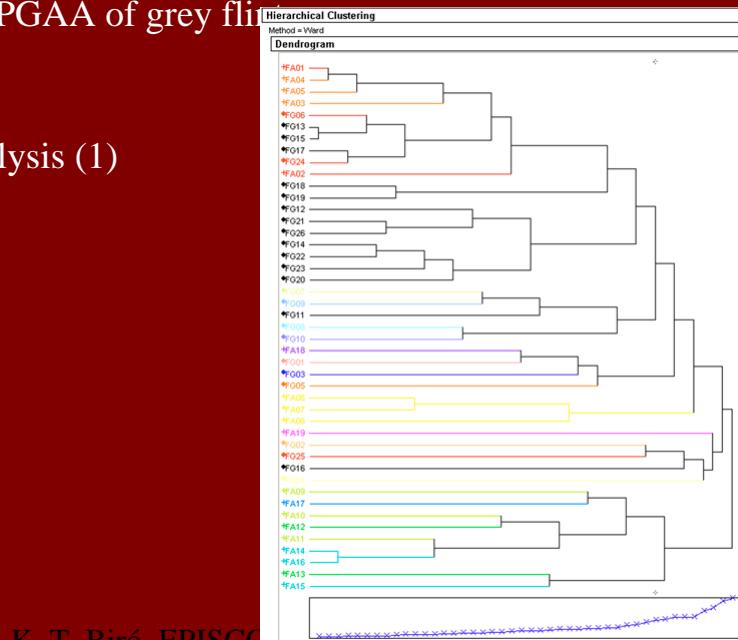
## Case study: PGAA of grey flint

Ternary plot  
(obsidian)



## Case study: PGAA of grey flint

Cluster analysis (1)



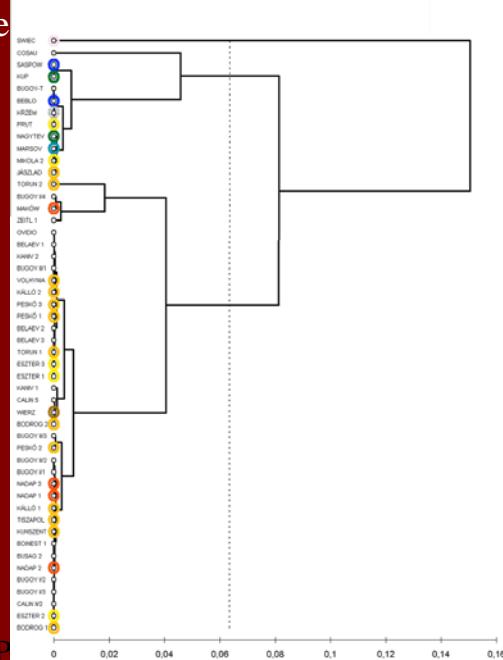
K. T. Biró, EPISCC

## Case study: PGAA of grey flint

### Cluster analysis (2)

K. T. Biró, EP

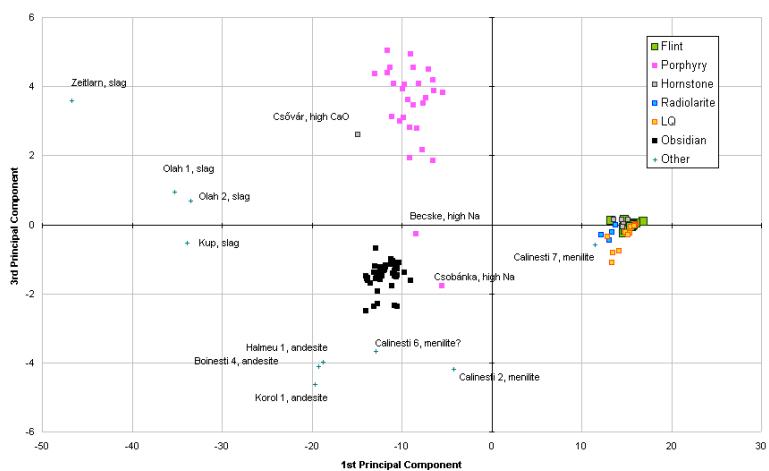
Dendrogram - Grey flint total



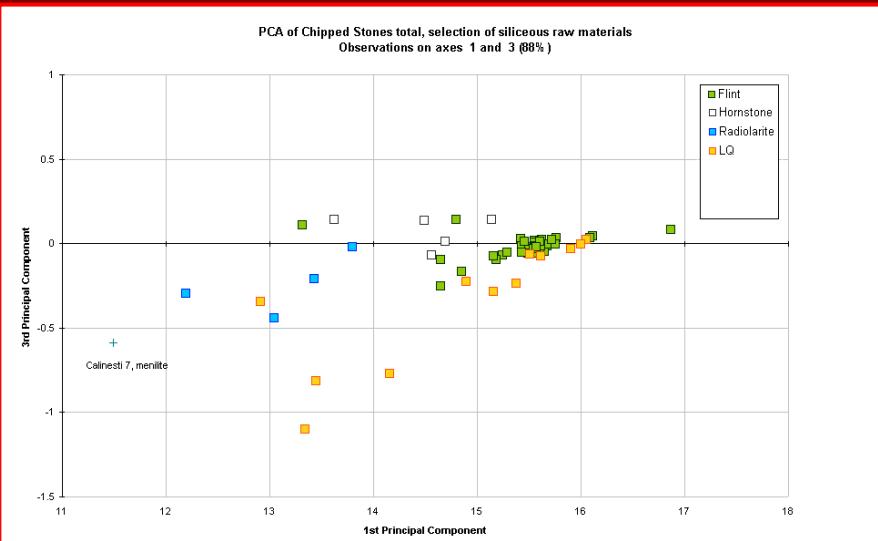
## Case study: PGAA of grey flint

K. T. Biró, EPISCON course 24/01/2007

PCA of Chipped Stones total  
Observations on axes 1 and 3 (88% )

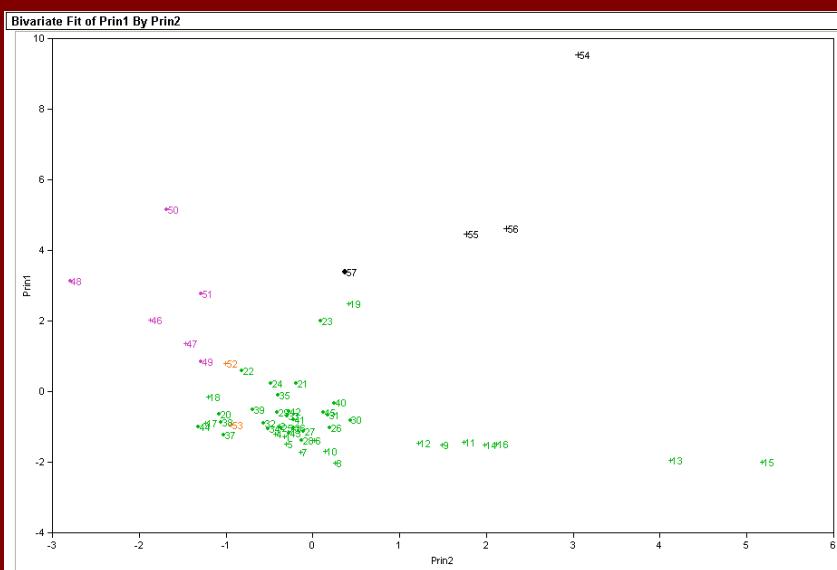


## Case study: PGAA of grey flint



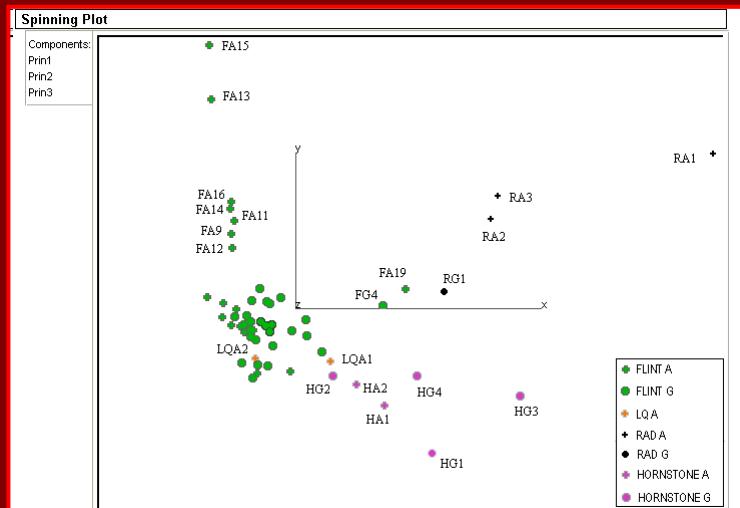
K. T. Biró, EPISCON course 24/01/2007

## Case study: PGAA of grey flint



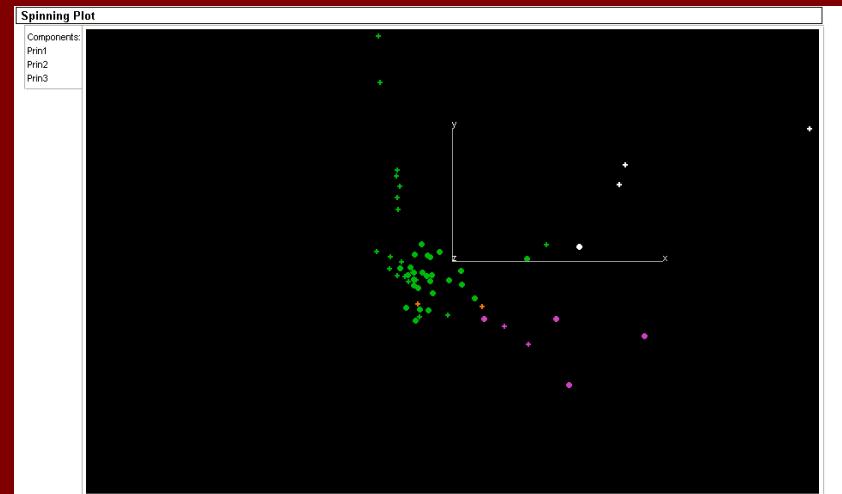
K. T. Biró, EPISCON course 24/01/2007

## Case study: PGAA of grey flint



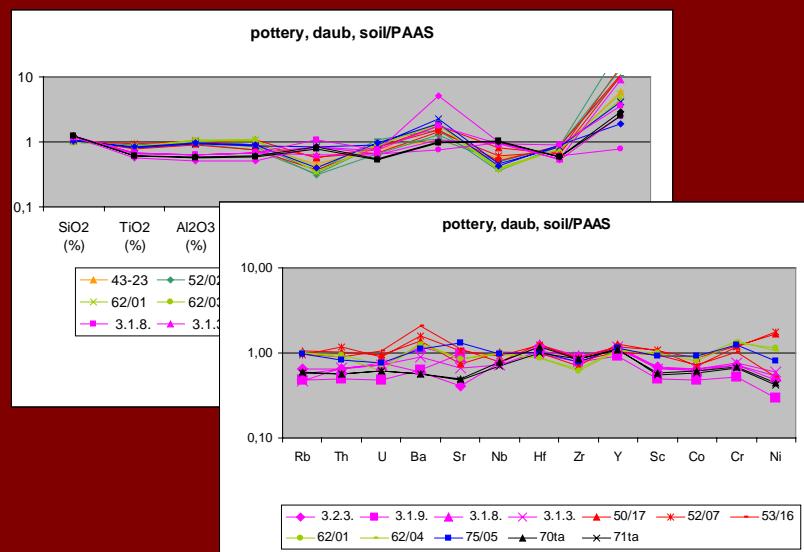
K. T. Biró, EPISCON course 24/01/2007

## Case study: PGAA of grey flint



K. T. Biró, EPISCON course 24/01/2007

## Other graphs



K. T. Biró, EPISCON course 24/01/2007