

Многомерное шкалирование

Стат. методы в
психологии
(Радчикова Н.П.)



Многомерное шкалирование

Многомерное шкалирование (МШ)
можно рассматривать как
альтернативу факторному анализу.





Многомерное шкалирование

Целью является поиск и интерпретация «латентных» (т.е. непосредственно не наблюдаемых) переменных, дающих возможность пользователю объяснить сходства между объектами, заданными точками в исходном пространстве признаков.



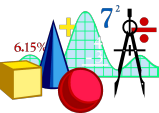


Многомерное шкалирование

Основная цель:

Выявление структуры исследуемого
множества объектов

(структура – набор основных
факторов (шкал), по которым
различаются и могут быть описаны
объекты)





Многомерное шкалирование

Исходная информация для МШ –
данные о различии или близости
объектов





Многомерное шкалирование

В психологии исходными данными для МШ являются субъективные суждения испытуемых о различии или сходстве стимулов (объектов).





Многомерное шкалирование

Считается, что в основе таких суждений лежит ограниченное количество субъективных признаков (критериев), определяющих различение стимулов, и человек, вынося свои суждения, явно или неявно учитывает эти критерии.

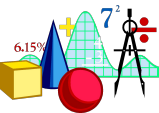




Многомерное шкалирование

Основная задача МШ:

Реконструкция психологического пространства, заданного небольшим числом измерений-шкал, и расположение в нем точек-стимулов таким образом, чтобы расстояние между ними наилучшим образом соответствовало исходным субъективным различиям.

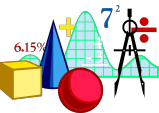




Многомерное шкалирование

Поэтому

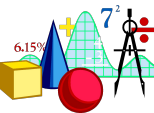
шкала в МШ интерпретируется как критерий, лежащий в основе различий СТИМУЛОВ.





Основная идея

Пример: пусть имеется матрица попарных расстояний (т.е. сходства некоторых признаков) между крупными белорусскими городами. Анализируя матрицу, стремятся расположить точки с координатами городов в двумерном пространстве (на плоскости), максимально сохранив реальные расстояния между НИМИ.





Основная идея

В общем случае метод МНШ позволяет таким образом расположить "объекты" (города в нашем примере) в пространстве некоторой небольшой размерности (в данном случае она равна двум), чтобы достаточно адекватно воспроизвести наблюдаемые расстояния между ними.



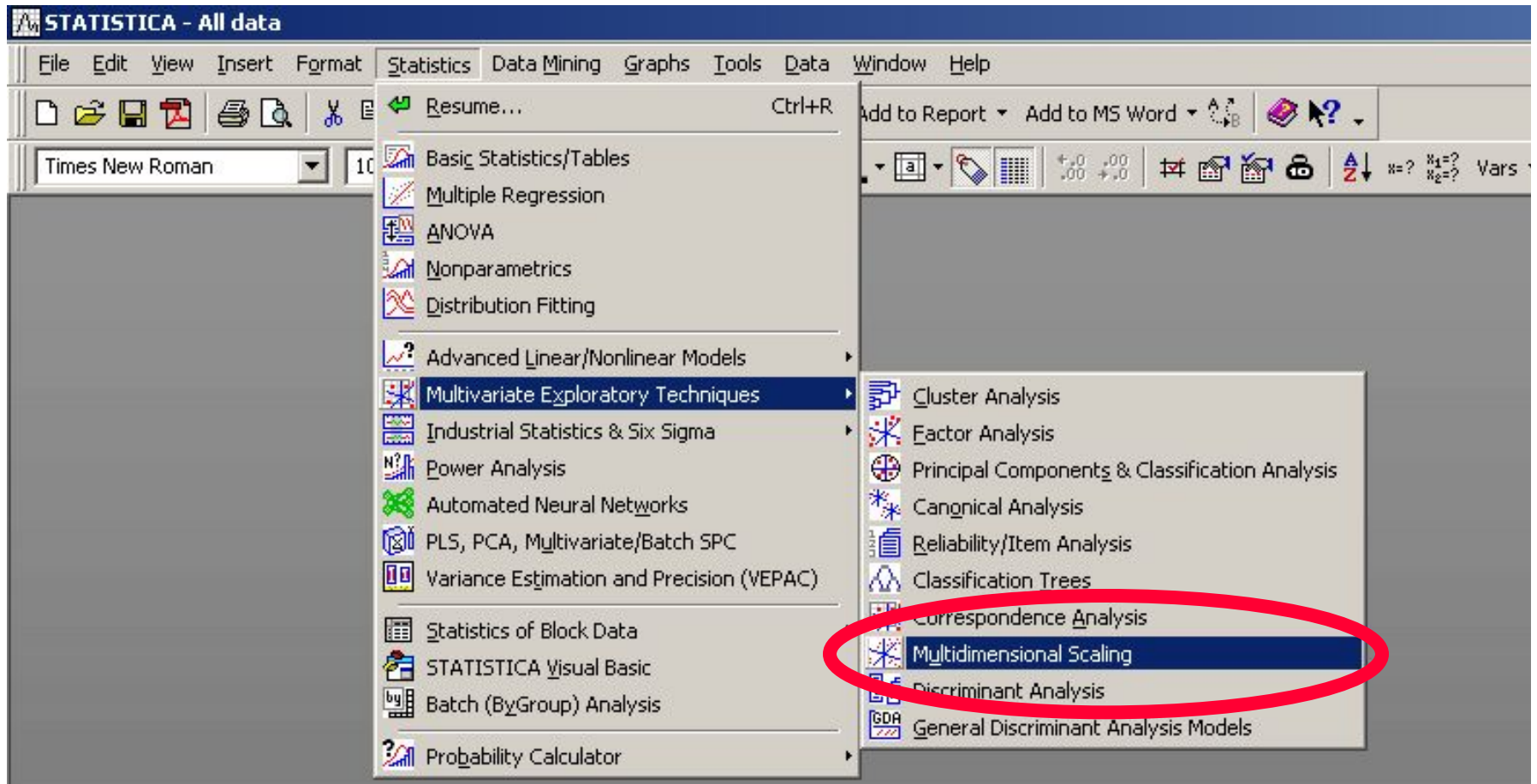


Основная идея

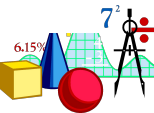
В результате можно "измерить" эти расстояния в терминах найденных латентных переменных. Так, в нашем примере можно объяснить расстояния в терминах пары географических координат Север/Юг и Восток/Запад.



Многомерное шкалирование



**Statistics ⇒ Multivariate Exploratory Techniques ⇒
Multidimensional Scaling**





Представление данных

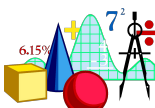
STATISTICA: Multidimensional Scaling

File Edit View Analysis Graphs Options Window Help

3, [Icons] Vars Cases ABC [Icons] [Icons] [Icons] [Icons] [Icons]

Data: Belarus.STA 10v * 14c

NUMERIC VALUES	1 MINSK	2 ORSHA	3 VITEBSK	4 MOGILEV	5 GOMEL	6 MOZIR	7 PINSK	8 BREST	9 GRODNO	10 POLOTZK
MINSK	0,00	8,0	8,5	7,0	11,0	9,0	9,0	13,0	9,7	7,5
ORSHA	8,00	0,0	3,0	2,7	9,2	11,3	15,7	20,5	17,0	6,0
VITEBSK	8,50	3,0	0,0	5,7	12,2	14,0	17,2	21,7	17,4	3,7
MOGILEV	7,00	2,7	5,7	0,0	6,5	8,2	13,7	19,3	17,0	8,0
GOMEL	11,00	9,2	12,2	6,5	0,0	5,0	13,3	19,6	19,6	14,5
MOZIR	9,00	11,3	14,0	8,2	5,0	0,0	8,5	15,0	16,0	15,2
PINSK	9,00	15,7	17,2	13,7	13,3	8,5	0,0	6,3	9,0	16,4
BREST	13,00	20,5	21,7	19,3	19,6	15,0	6,3	0,0	7,0	20,0
GRODNO	9,70	17,0	17,4	17,0	19,6	16,0	9,0	7,0	0,0	15,0
POLOTZK	7,50	6,0	3,7	8,0	14,5	15,2	16,4	20,0	15,0	0,0
Mean	8,00	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Std. Dev.	0,00	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
No. Cases	10,00									
Matrix	3,00									





Представление данных

Матрица может представлять

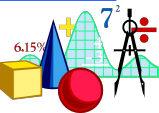
1 – корреляцию

2 – сходство

3 – различие

4 – ковариацию

ЭТИ ОБОЗНАЧЕНИЯ ЯВЛЯЮТСЯ СТАНДАРТНЫМИ!



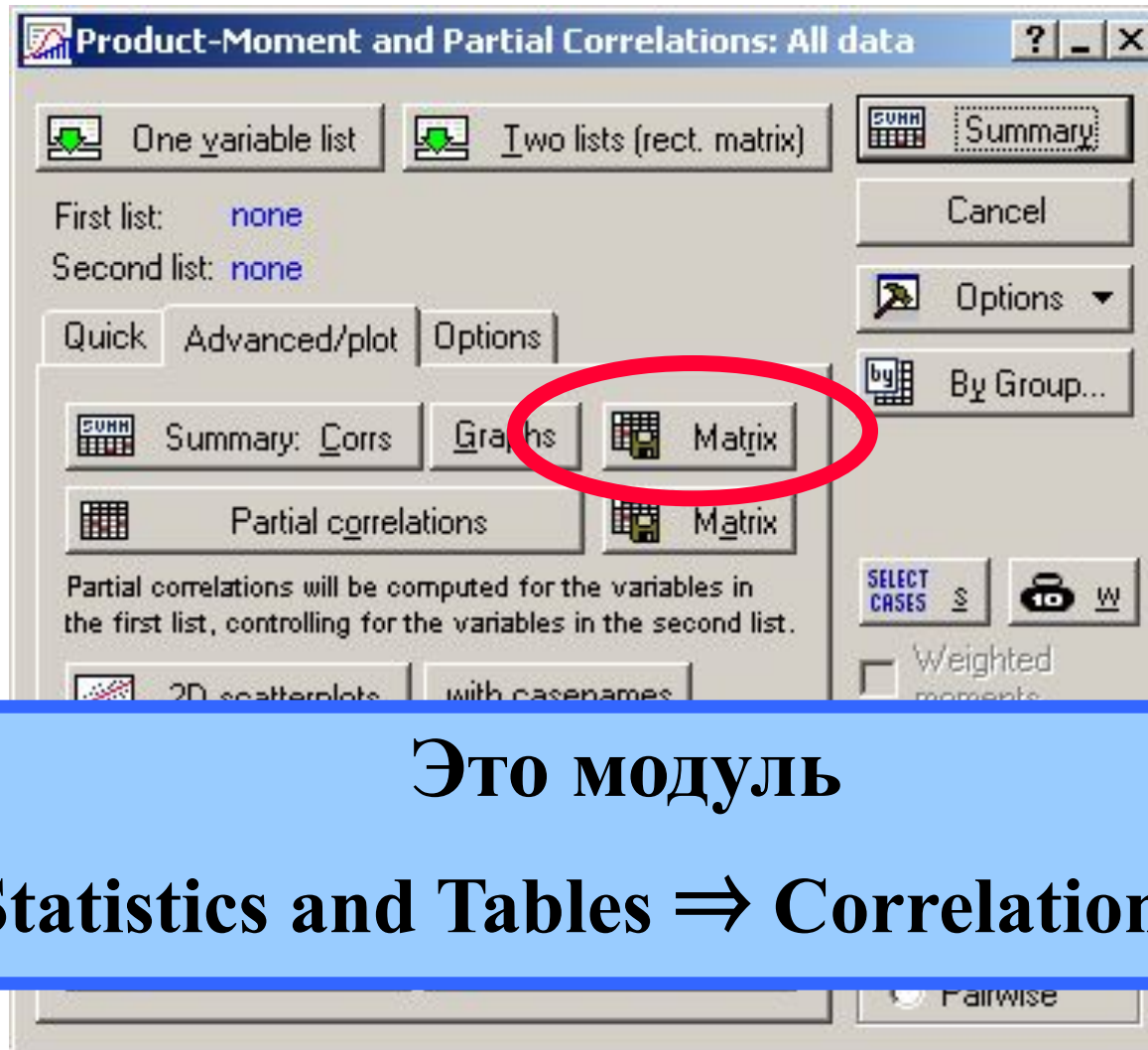


Как получить матрицу?

- 1) В модуле, где считают корреляцию
- 2) В модуле кластерного анализа
- 3) Самим посчитать

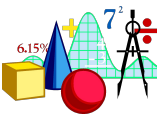


Как получить матрицу?



Это модуль

Basic Statistics and Tables \Rightarrow Correlation Matrices

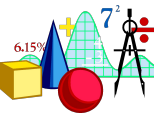
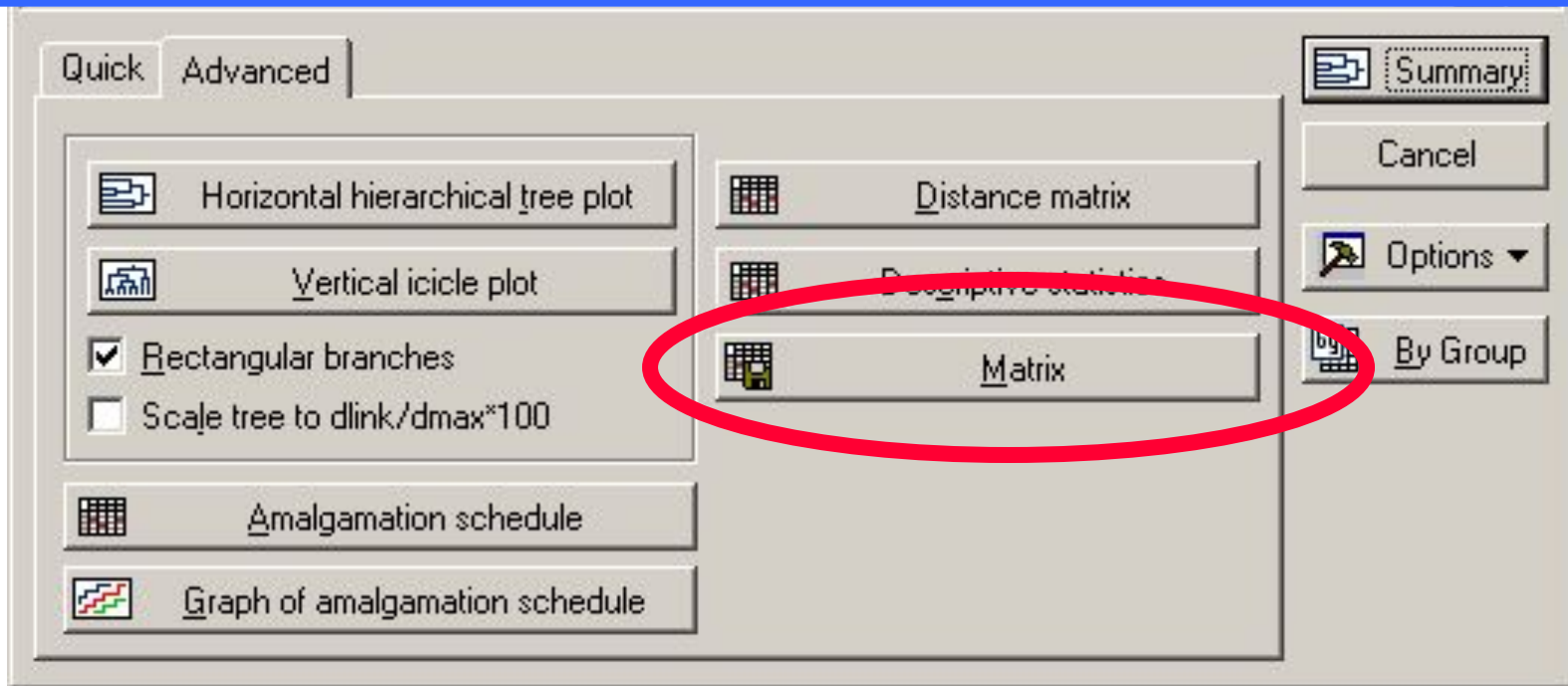


Как получить матрицу?



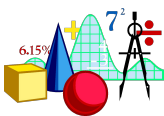
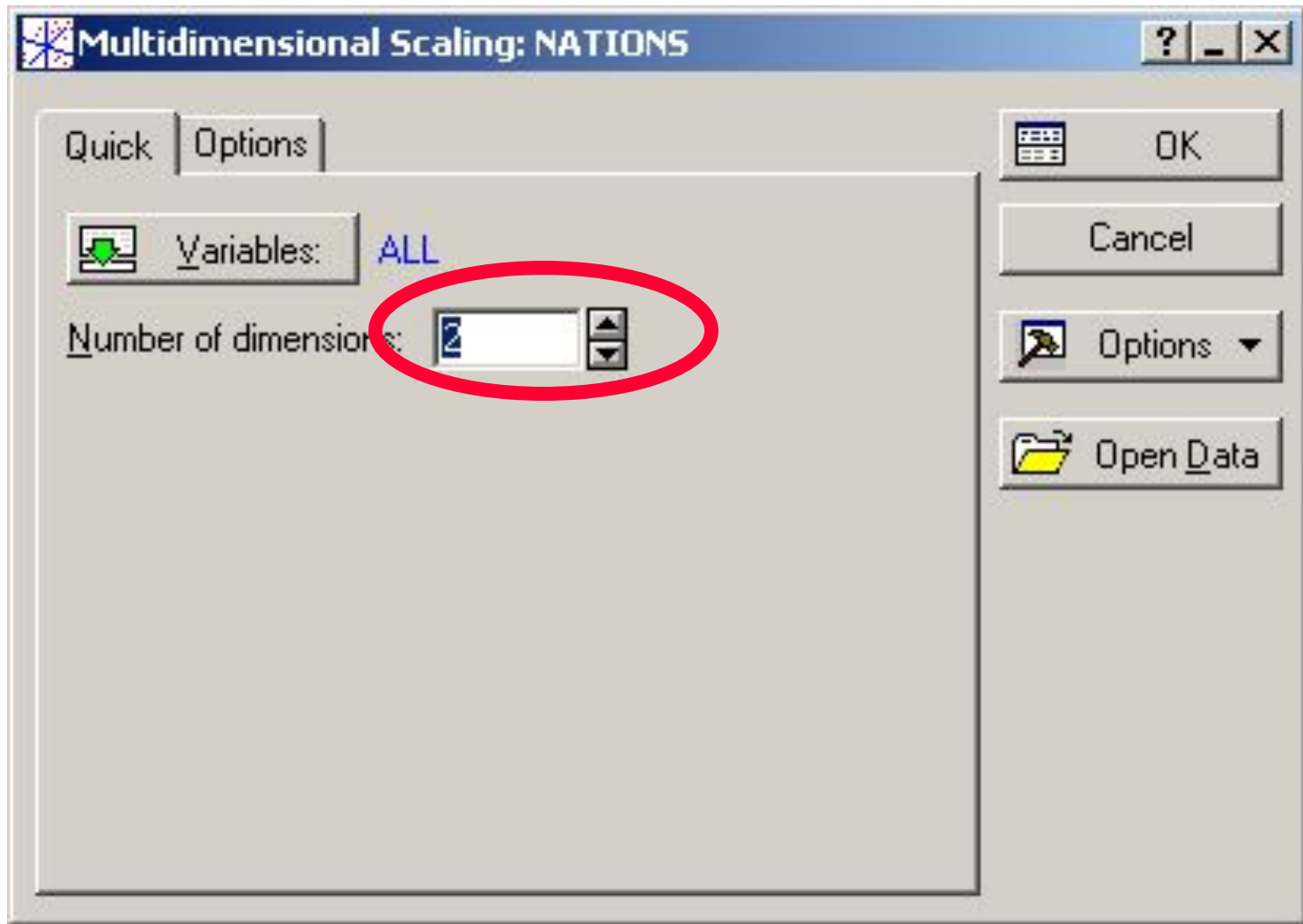
Это модуль

Cluster Analysis \Rightarrow Joining (Tree Clustering)





Многомерное шкалирование



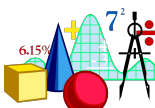


Результаты

Parameter Estimation							
iter.	[dim=2]	D-star		D-hat		d-hat	
s: t:	cosin	step	raw stress	alienation	raw stress	stress	
59	1	,965	,670			,0000000	,0000127
60	1	,940	,596			,0000000	,0000112
61	1	,920	,534			,0000000	,0000101
62	1	,933	,539			,0000000	,0000090
63	1	,955	,586			,0000000	,0000080
64	1	,962	,619			,0000000	,0000071
65	1	,957	,619			,0000000	,0000063
66	1	,947	,595			,0000000	,0000055
67	1	,936	,565			,0000000	,0000050
66	*			,0000000	,0000098	,0000000	,0000050
50	0					,0000002	,0000414
51	1		,100			,0000002	,0000405
52	1	1,000	,399			,0000001	,0000370
53	1	,994	,620			,0000001	,0000321
54	1	,997	,725			,0000001	,0000272
55	1	,987	,734			,0000001	,0000232
56	1	,991	,748			,0000000	,0000198
57	1	,995	,765			,0000000	,0000168
58	1	,976	,716			,0000000	,0000145

Estimation procedure converged

Cancel OK





Результаты

Results: NATIONS [?] [-] [X]

12 vars. from file
Number of dimensions: 2
Start config.: (last final)
Last iteration computed: 35; Best iteration: 20
D-star: Raw stress = 8,228679; Alienation = ,2373336
D-hat: Raw stress = 5,171562; Stress = ,1895089

Quick | Advanced | Review & save

Summary [Final configuration]

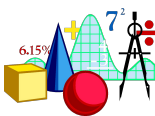
Graph final configuration, 2D

Graph final configuration, 3D

Shepard diagram

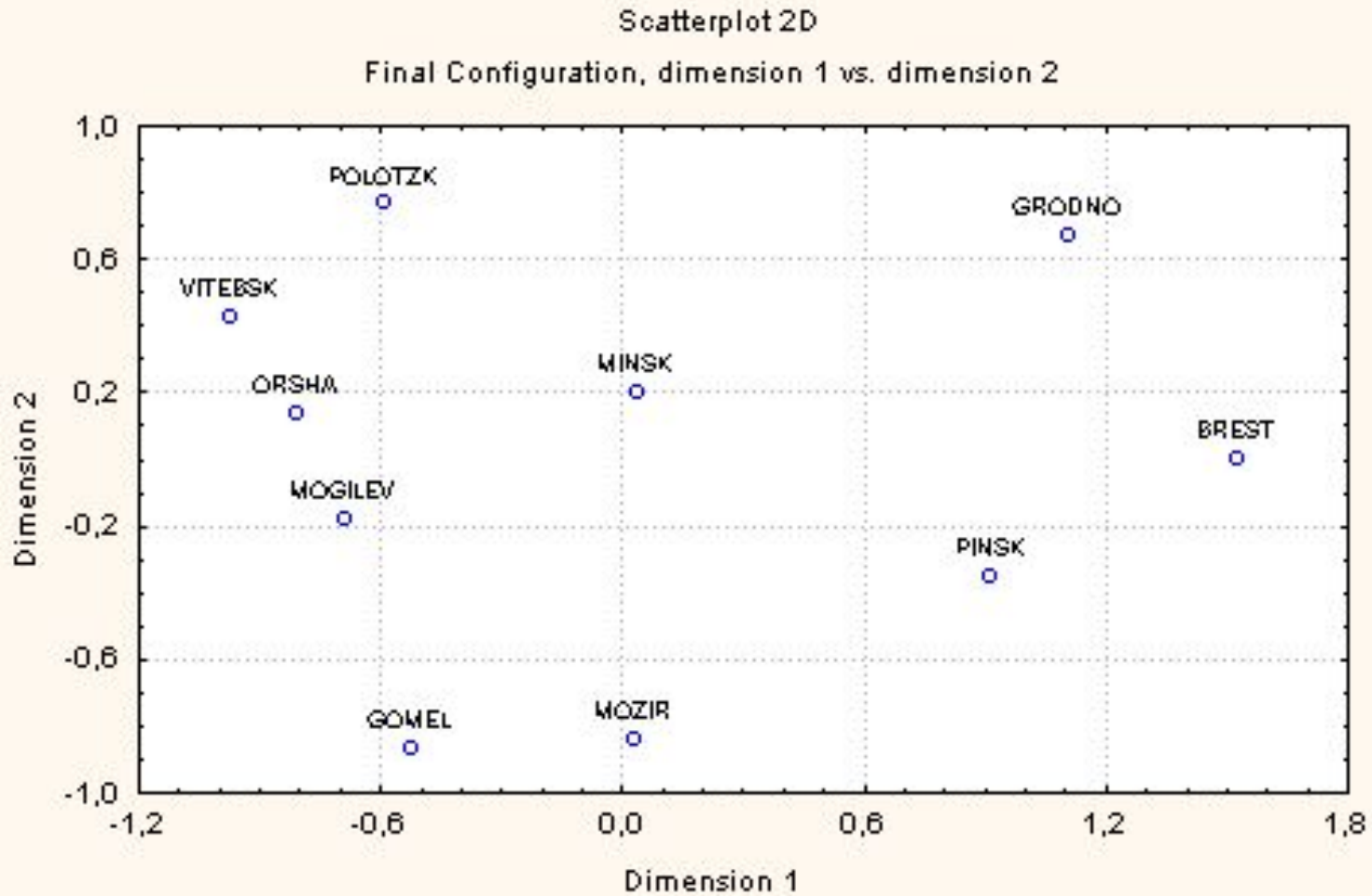
Cancel

Options ▾





Результаты





Ориентация осей

**Ориентация осей может быть
выбрана произвольной!**





Результаты

Results: NATIONS [?] [-] [X]

```
12 vars. from file
Number of dimensions: 2
Start config.: (last final)
Last iteration computed: 35; Best iteration: 20
D-star: Raw stress = 8,228679; Alienation = ,2373336
D-hat: Raw stress = 5,171562; Stress = ,1895089
```

Quick | Advanced | Review & save

Summary: Final configuration

Graph final configuration, 2D

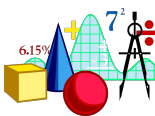
Graph final configuration, 3D

Shephard diagram

Summary

Cancel

Options ▾





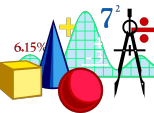
Результаты

Final Configuration (belarus.sta)

MULTIDIM SCALING

D-star: Raw stress = ,0000000; Alienation = ,0000098
D-hat: Raw stress = ,0000000; Stress = ,0000050

	DIM. 1	DIM. 2
MINSK	,031902	,202300
ORSHA	-,810804	,138991
VITEBSK	-,978265	,430584
MOGILEV	-,694163	-,174564
GOMEL	-,527156	-,860427
MOZIR	,025860	-,836591
PINSK	,913378	-,347500
BREST	1,526213	,002790
GRODNO	1,106199	,673348
POLOTZK	-,593164	,771069





Результаты

Results: NATIONS [?] [-] [X]

```
12 vars. from file
Number of dimensions: 2
Start config.: (last final)|
Last iteration computed: 25: Best iteration: 29
D-star: Raw stress =
D-hat: Raw stress =
```

Оценки расстояний

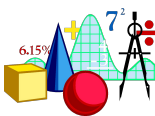
Quick | **Advanced** | Preview & save

Summary Summary

Cancel

Options ▾

Summary: Final configuration	Graph final configuration, 2D
D-hat values	Graph final configuration, 3D
D-star values	Graph D-hat vs. distances
Distance matrix	Graph D-star vs. distances
Summary statistics	Shepard diagram





Результаты

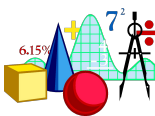
Results: NATIONS

```
12 vars. from file
Number of dimensions: 2
Start config.: (last f
Last iteration computed
D-star: Raw stress = 8
D-hat: Raw stress = 5
```

Воспроизведенная матрица расстояний

Quick | **Advanced** | Preview & save

Summary: Final configuration	Graph final configuration, 2D	Summary Cancel Options ▾
D-hat values	Graph final configuration, 3D	
D-star values	Graph D-hat vs. distances	
Distance matrix	Graph D-star vs. distances	
Summary statistics	Shepard diagram	





Результаты

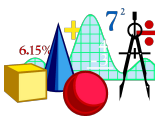
Results: NATIONS

```
12 vars. from file
Number of dimensions: 2
Start config.: (last fi
Last iteration computed
D-star: Raw stress = 8,
D-hat: Raw stress = 5,
```

Quick | **Advanced** | Preview & save

Summary: Final configuration	Graph final configuration, 2D	Cancel
D-hat values	Graph final configuration, 3D	Options ▾
D-star values	Graph D-hat vs. distances	
Distance matrix	Graph D-star vs. distances	
Summary statistics	Shepard diagram	

**Итоговая статистика
по предыдущим трем
кнопкам**





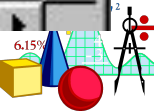
Результаты

Final Configuration (belarus.sta)

MULTIDIM SCALING

D-star: Raw stress = ,0000000; Alienation = ,0000097
D-hat: Raw stress = ,0000000; Stress = ,0000049

	Distance	D-star	D-hat
D(4, 2)	,325345	,325345	,325345
D(3, 2)	,326357	,326357	,326357
D(10, 3)	,511204	,511204	,511204
D(6, 5)	,553621	,553621	,553621
D(4, 3)	,647492	,647472	,647482
D(10, 2)	,647472	,647492	,647482
D(8, 7)	,719436	,719436	,719436
D(5, 4)	,727284	,727284	,727284
D(9, 8)	,796003	,796003	,796003
D(4, 1)	,812487	,812487	,812487
D(10, 1)	,836342	,836316	,836329
D(2, 1)	,836316	,836342	,836329
D(10, 4)	,925187	,925187	,925187
D(6, 4)	,990687	,990687	,990687
D(7, 6)	,999303	,999303	,999303





Результаты

Оценка качества модели

Results: NATIONS

Iteration: 20
Function = ,2373336
D-hat: Raw stress = 5,171562; Stress = ,1895089

Quick | Advanced | Review & save

Summary

Summary: Final configuration | Graph final configuration, 2D

D-hat values | Graph final configuration, 3D

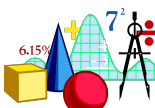
D-star values | Graph D-hat vs. distances

Distance matrix | Graph D-star vs. distances

Summary statistics | Shepard diagram

Cancel

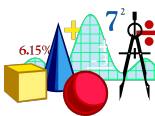
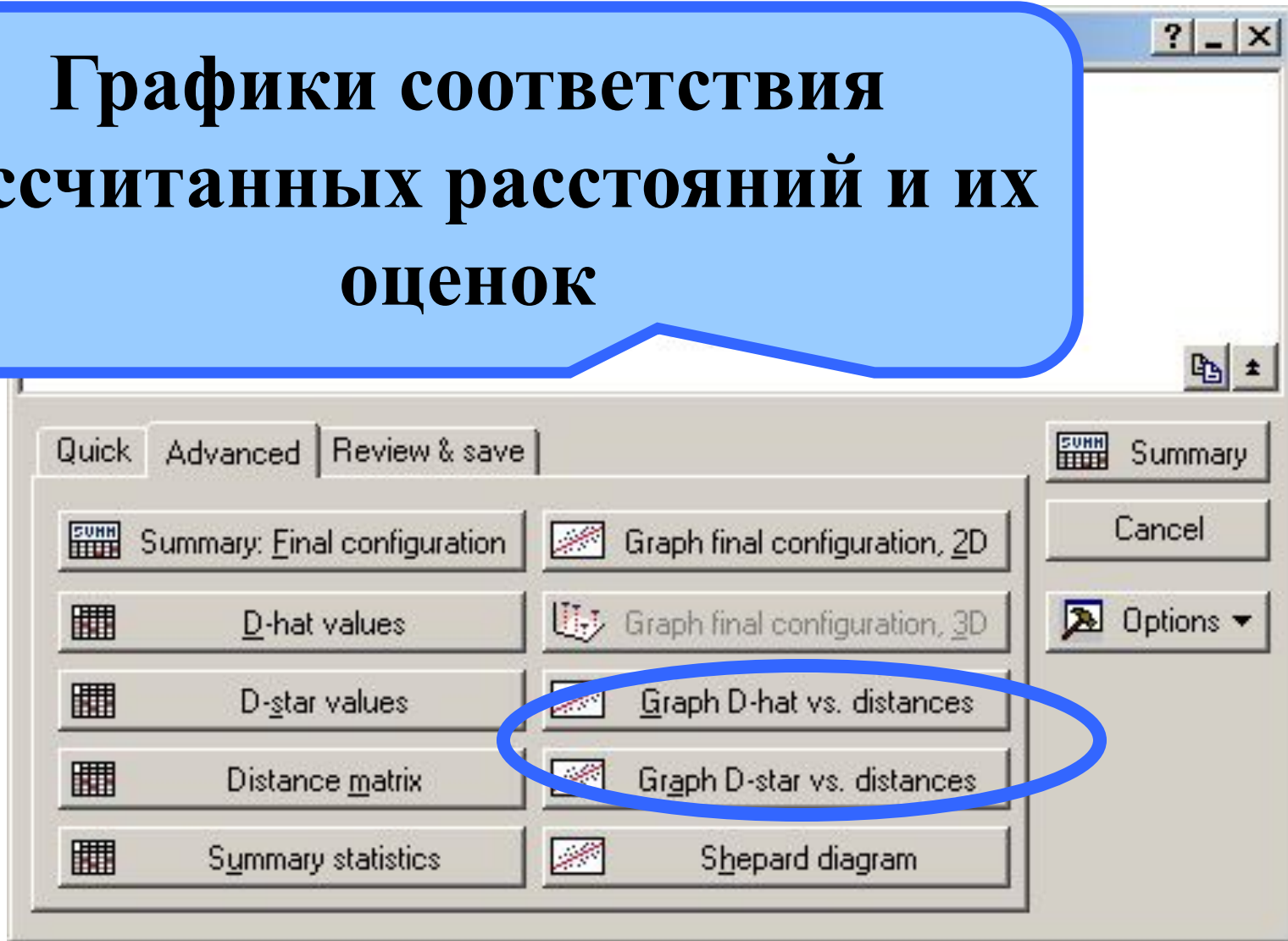
Options





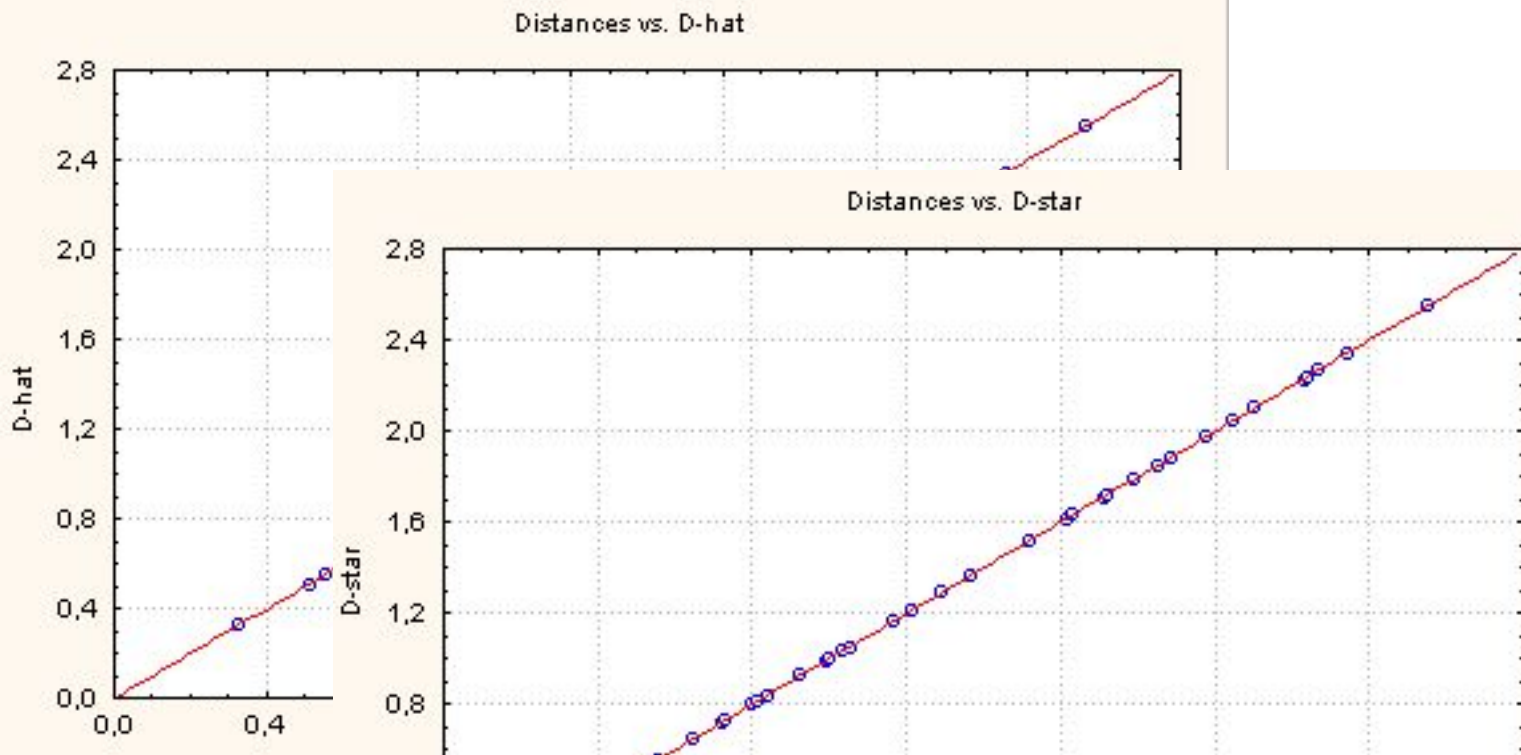
Результаты

Графики соответствия рассчитанных расстояний и их оценок

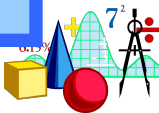




Графики



Точки должны лежать как можно ближе к диагонали





Результаты

Диаграмма Шепарда

Results: NATIONS

12 vars. from file
Number of dimensions: 2
Start config.: (last final)

Iteration: 20
on = ,2373336
= ,1895089

Quick | Advanced | Review & save

Summary

Summary: Final configuration | Graph final configuration, 2D

D-hat values | Graph final configuration, 3D

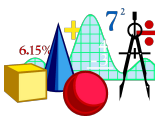
D-star values | Graph D-hat vs. distances

Distance matrix | Graph D-star vs. distances

Summary statistics | **Shepard diagram**

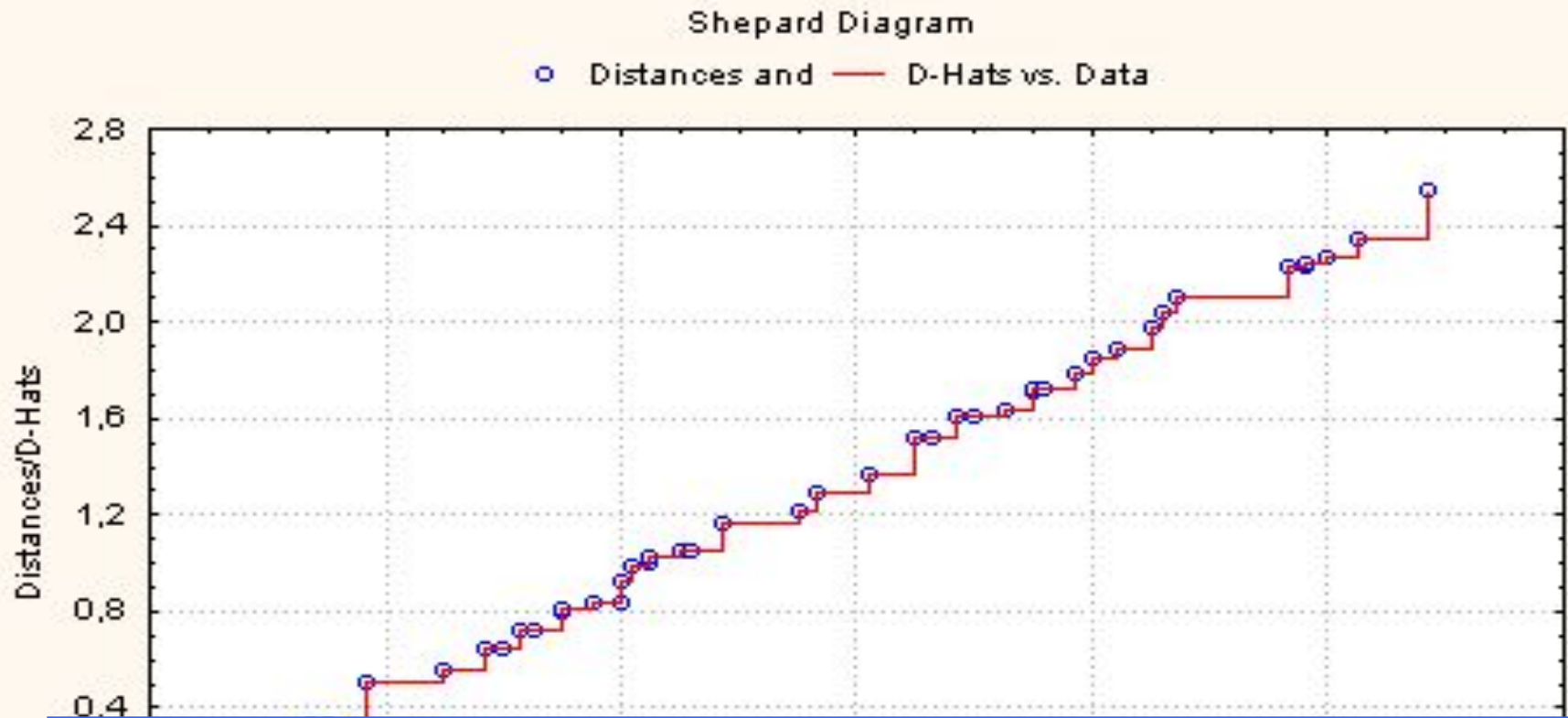
Cancel

Options





Результаты



Точки должны лежать как можно ближе к ступенчатой ломаной





Пример 2

**В файле данных
приведены
сведения о
доходах в 1991
году**

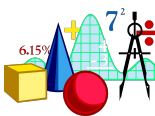
STATISTICA: Multidimensional Scaling

File Edit View Analysis Graphs Options Window Help

0, [Icons] Vars Cases [Icons]

Data: NATIONS.STA 12v * 16c

NUMERIC VALUES	1 BRAZIL	2 CONGO	3 CUBA	4 EGYPT	5 FRANCE	6 INDIA	7 ISRAEL	8 JAPAN	9 CHINA	10 RUSSIA	11 USA	12 YUGOSLAV	13 Means	14 Std.Dev.
BRAZIL	0,00	4,83	5,28	3,44	4,72	4,50	3,83	3,50	2,39	3,06	5,39	3,17	0,00	0,00
CONGO	4,83	0,00	4,56	5,00	4,00	4,83	3,33	3,39	4,00	3,39	2,39	3,50	0,00	0,00
CUBA	5,28	4,56	0,00	5,17	4,11	4,00	3,61	2,94	5,50	5,44	3,17	5,11	0,00	0,00
EGYPT	3,44	5,00	5,17	0,00	4,78	5,83	4,67	3,83	4,39	4,39	4,28	4,28	0,00	0,00
FRANCE	4,72	4,00	4,11	4,78	0,00	3,44	4,00	4,22	3,67	5,06	4,28	4,72	0,00	0,00
INDIA	4,50	4,83	4,00	5,83	3,44	0,00	4,11	4,50	4,11	4,50	4,17	4,44	0,00	0,00
ISRAEL	3,83	3,33	3,61	4,67	4,00	4,11	0,00	4,83	3,00	4,17	5,94	4,28	0,00	0,00
JAPAN	3,50	3,39	2,94	3,83	4,22	4,50	4,83	0,00	4,11	4,17	6,06	4,28	0,00	0,00
CHINA	2,39	4,00	5,50	4,39	3,67	4,11	3,00	4,17	0,00	4,61	6,06	4,28	0,00	0,00
RUSSIA	3,06	3,39	5,44	4,39	5,06	4,50	4,17	4,61	4,61	0,00	6,06	4,28	0,00	0,00
USA	5,39	2,39	3,17	3,33	5,94	4,28	5,94	6,06	6,06	0,00	0,00	4,28	0,00	0,00
YUGOSLAV	3,17	3,50	5,11	4,28	4,72	4,00	4,44	4,28	4,28	4,28	0,00	0,00	0,00	0,00
Means	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Std.Dev.	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

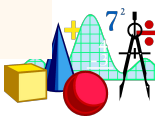
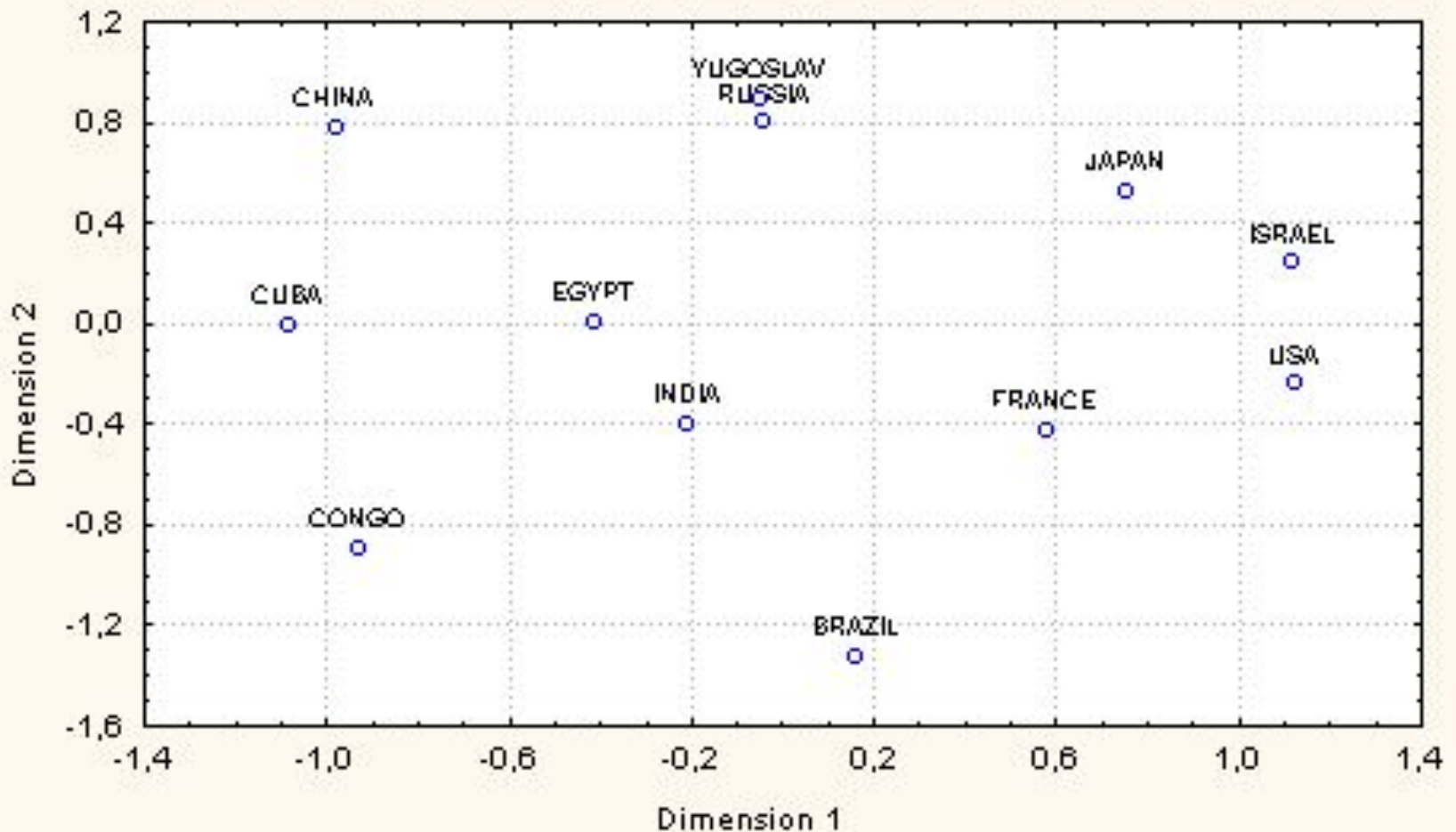




Пример 2

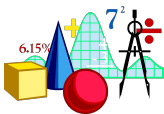
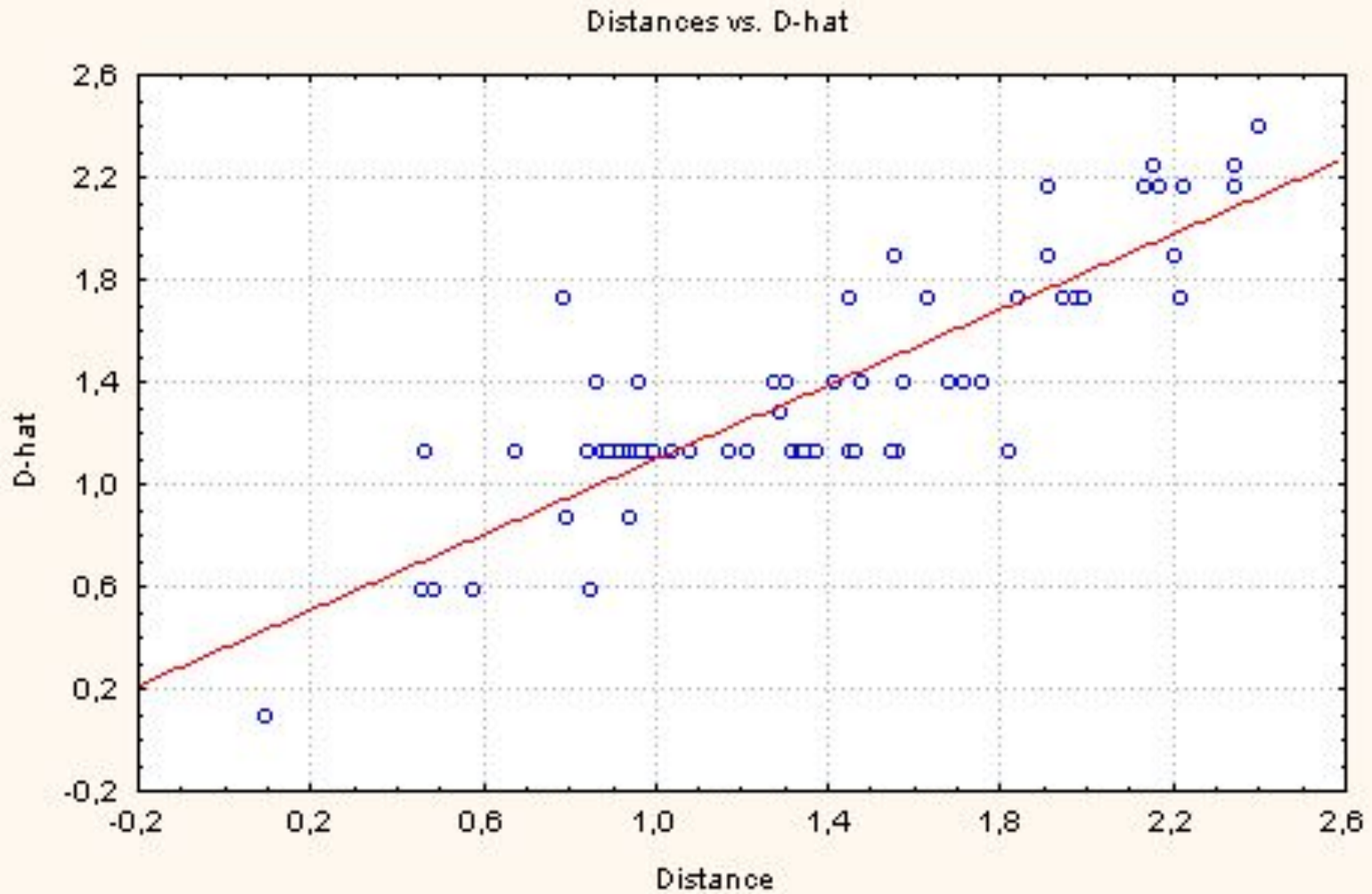
Scatterplot 2D

Final Configuration, dimension 1 vs. dimension 2



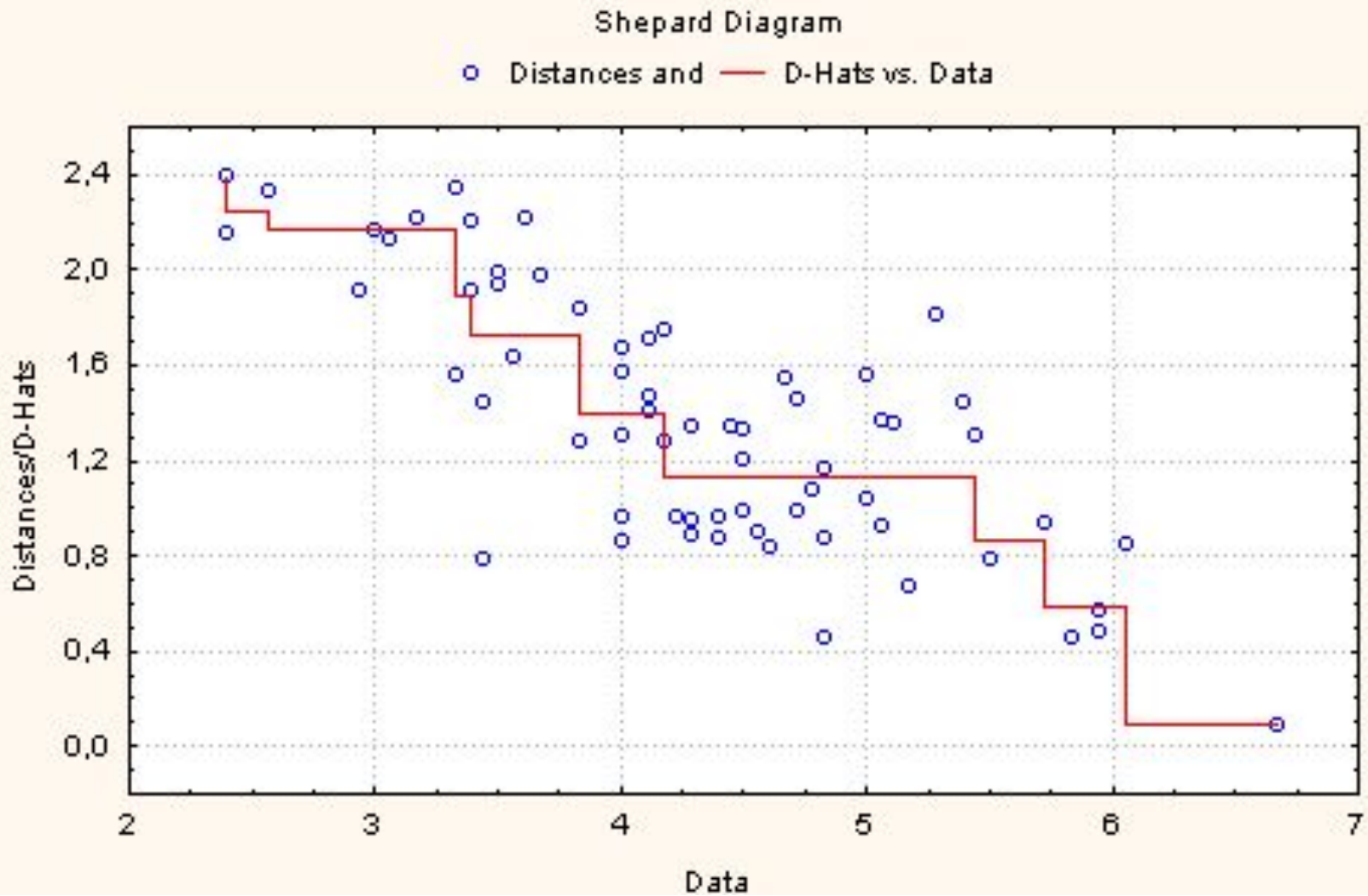


Пример 2





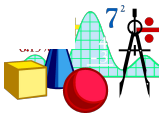
Пример 2





Пример 2

**А можно ли
попробовать
взять другое
количество
размерностей?**

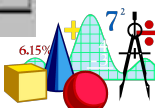




Пример 2

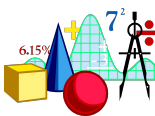
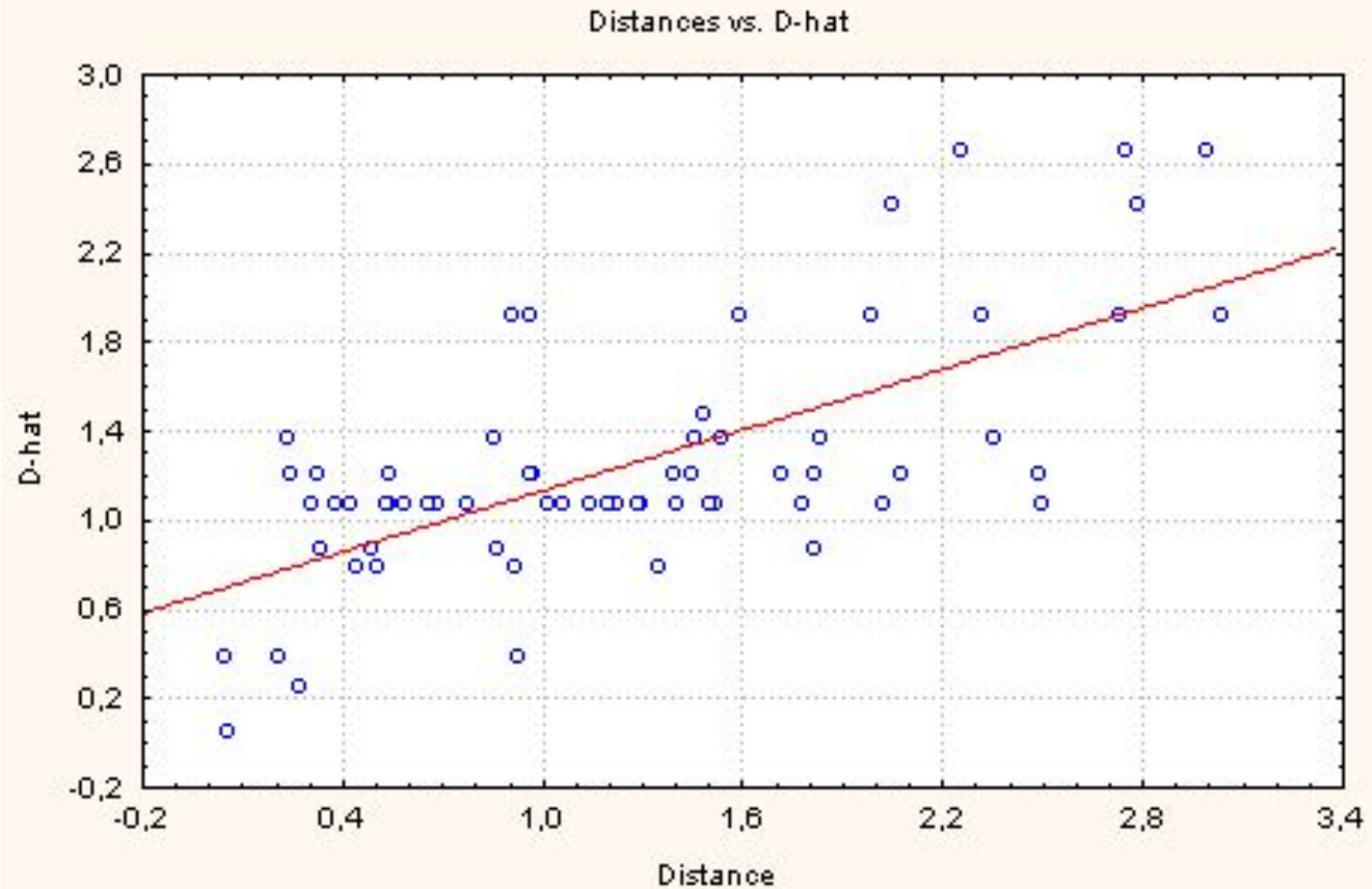
КОНЕЧНО! Берем одну шкалу:

Final Configuration (nations.sta)	
MULTIDIM SCALING	D-star: Raw stress = 33,52059; Alienation = ,4682254 D-hat: Raw stress = 22,32512; Stress = ,3937455
	DIM. 1
BRAZIL	,86083
CONGO	-1,63220
CUBA	-,95388
EGYPT	-,62037
FRANCE	,43812
INDIA	-,41638
ISRAEL	1,39756
JAPAN	1,09520
CHINA	-1,38772
RUSSIA	-,04274
USA	1,35839
YUGOSLAV	-,09679



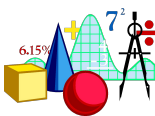
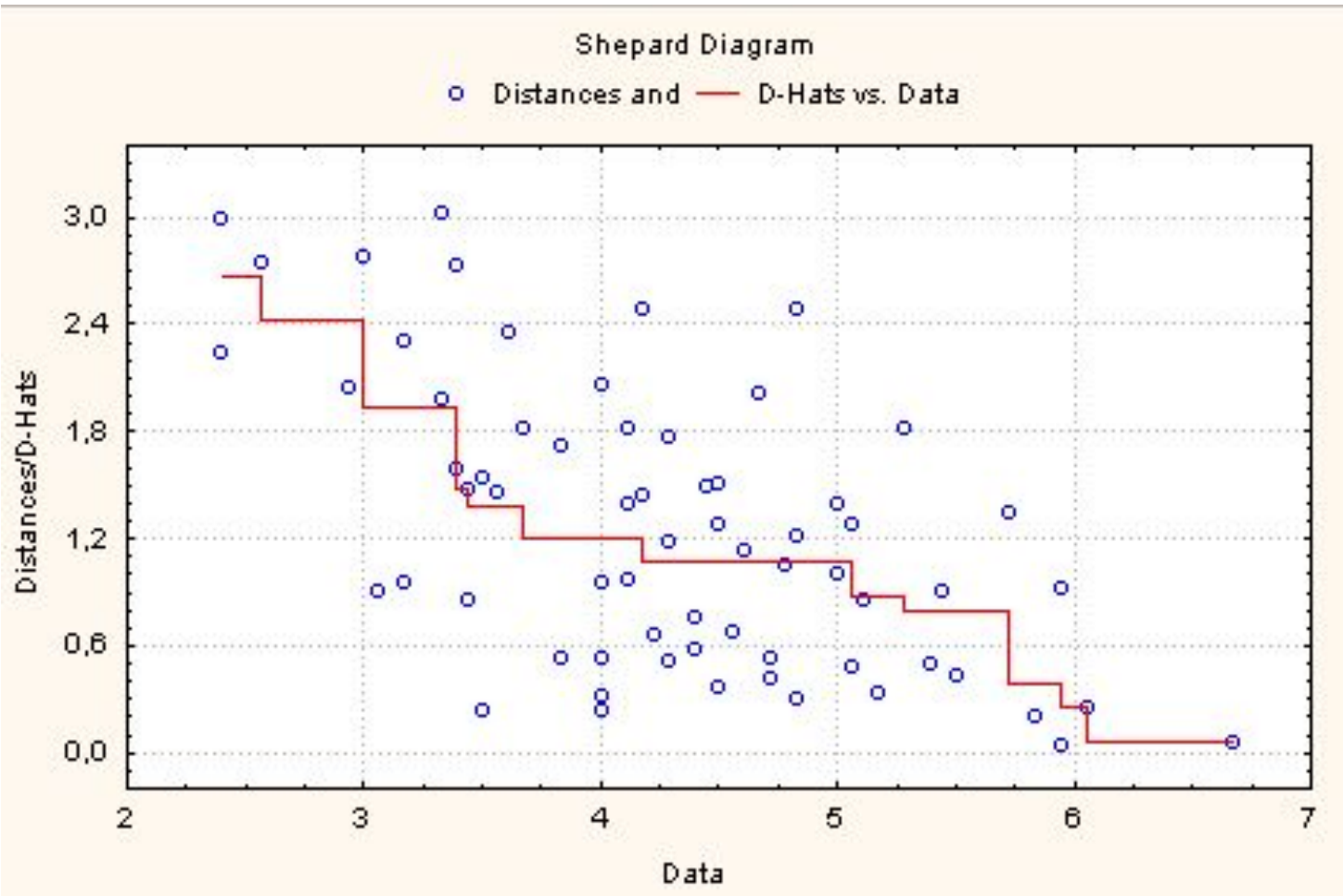


Пример 2





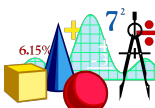
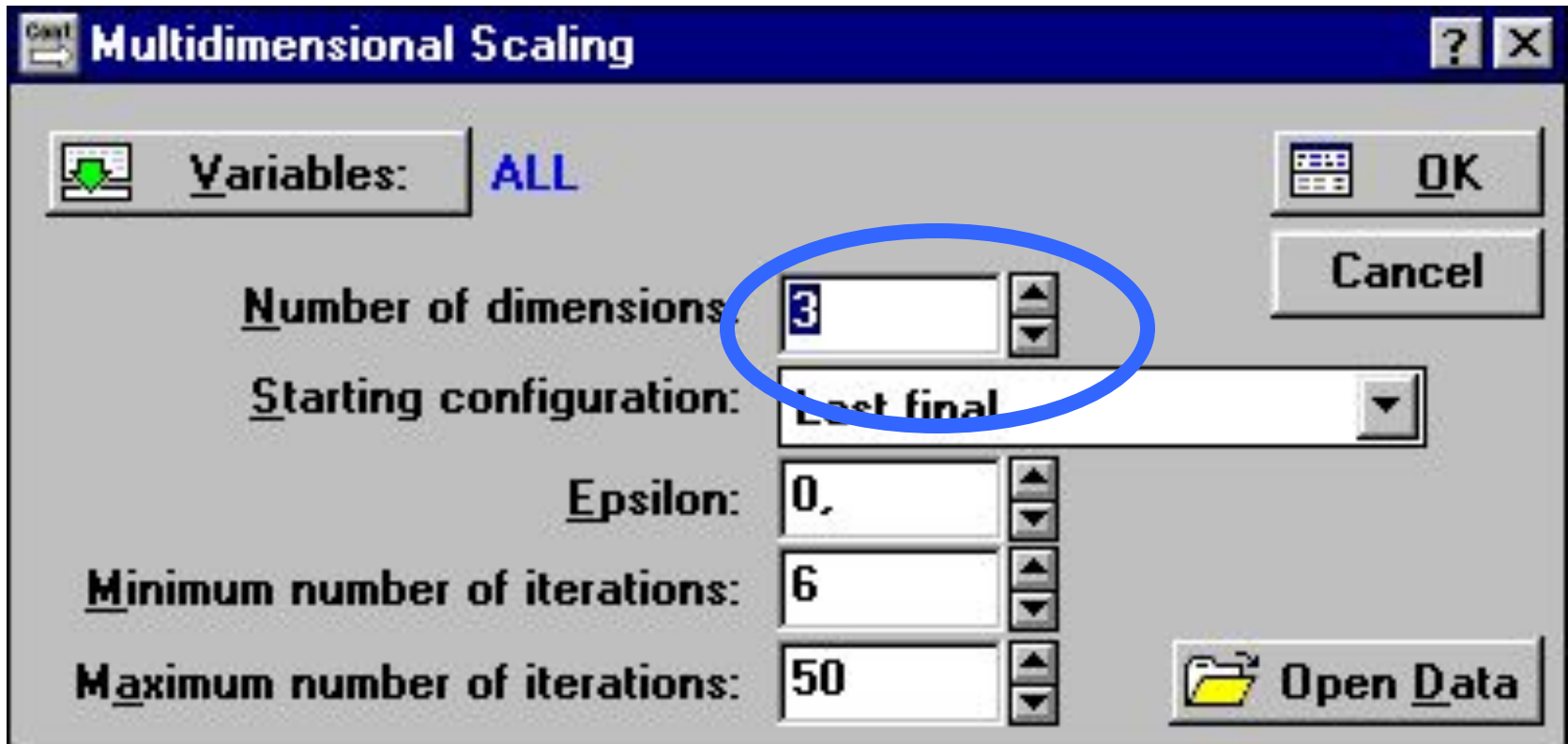
Пример 2





Пример 2

Берем три шкалы:

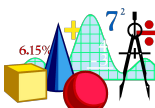
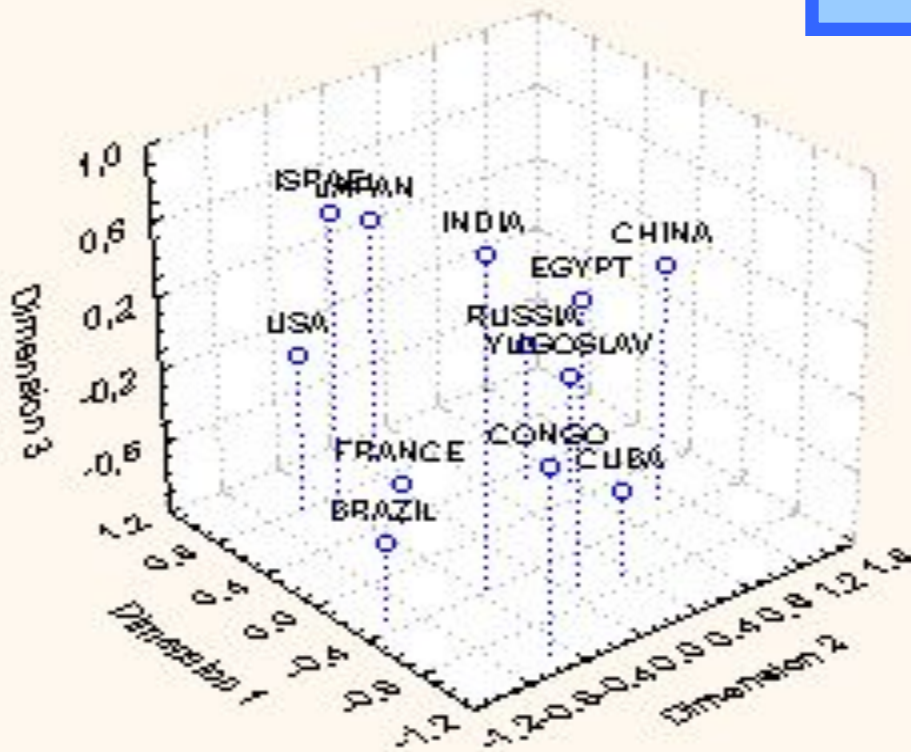




Пример 2

**Ну теперь-то
все понятно!**

Scatterplot 3D
Final Configuration
Dimension 2 vs. Dimension 1 vs. Dimension 3





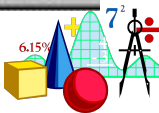
Пример 2

Вообще-то так понятнее:

Final Configuration (nations.sta)

MULTIDIM SCALING
D-star: Raw stress = 2,704252; Alienation = ,1367164
D-hat: Raw stress = 1,671670; Stress = ,1077442

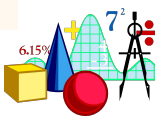
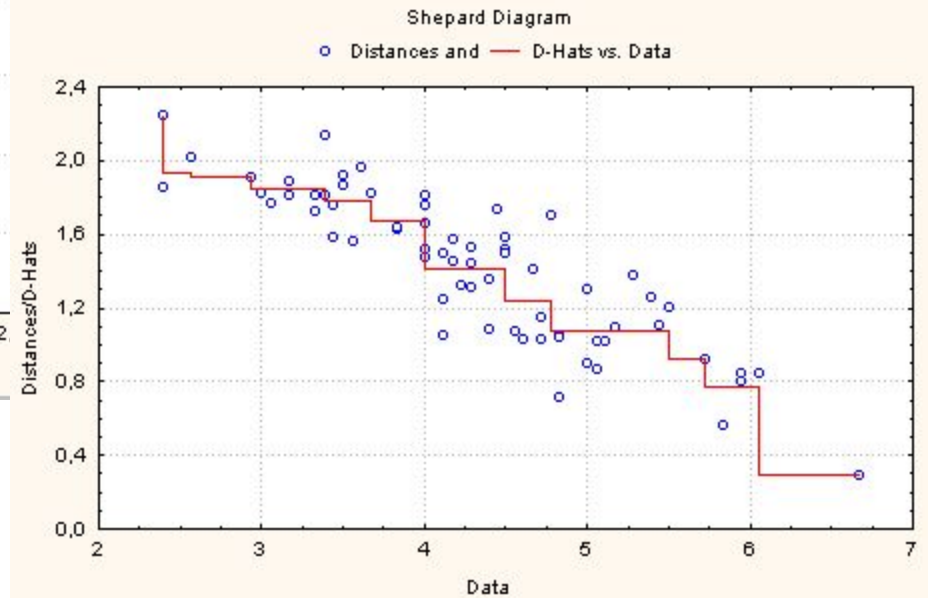
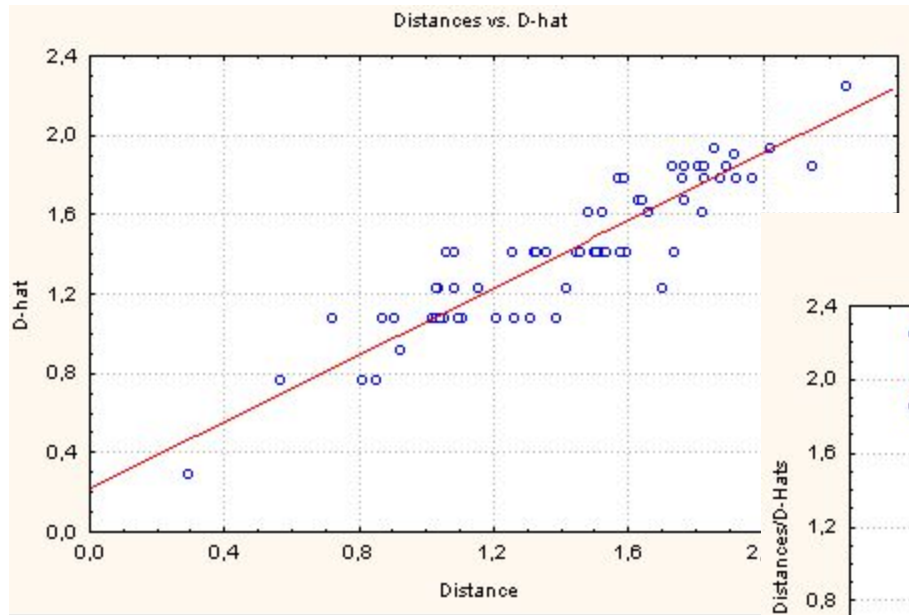
	DIM. 1	DIM. 2	DIM. 3
BRAZIL	-,30178	-,984157	-,569940
CONGO	-1,03957	-,529222	,027850
CUBA	-,70762	,336954	-,523638
EGYPT	-,63434	,096288	,541078
FRANCE	,37551	-,234263	-,785948
INDIA	-,37686	-,330311	,801502
ISRAEL	,66140	-,454461	,649067
JAPAN	,98086	,112078	,342146
CHINA	-,25729	1,085260	,306949
RUSSIA	,32158	,637774	-,249051
USA	,80806	-,570977	-,136612
YUGOSLAV	,17006	,835038	-,403403





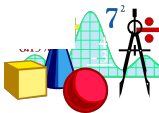
Пример 2

Оценим модель:





**А как вообще
выбрать
наилучшее
количество
размерностей?**





Выбор размерности пространства

- 1) **Критерий каменистой осыпи: на графике зависимости стресса от размерности**

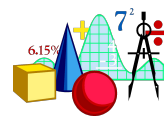


Выбор размерности пространства

Results

12 vars. from file D:\STAT5~1.5\EXAMPLES\NATIONS.STA
Number of dimensions: 2
Start config.: (last final)
Last iteration computed: 32; Best iteration: 13
D-star: Raw stress = 8,152610; Minimum = ,0062499
D-hat: Raw stress = 5,16613; Stress = ,1894094

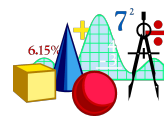
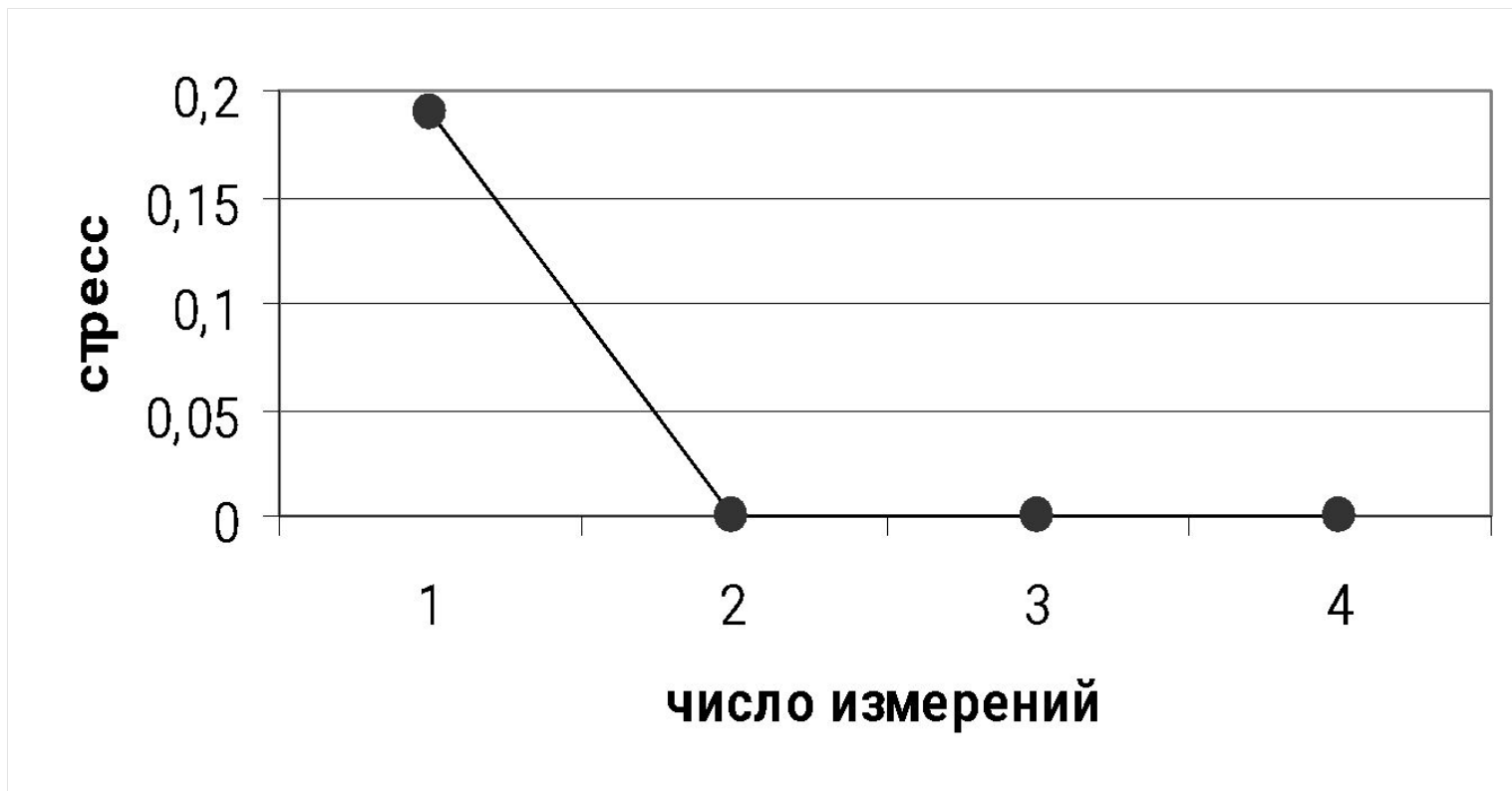
Final configuration Graph final configuration, 2D OK
D-hat values Graph final configuration, 3D Cancel
D-star values Graph D-hat vs. distances
Distance matrix Graph D-star vs. distances
Summary Shepard diagram
Save final configuration





Выбор размерности пространства

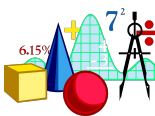
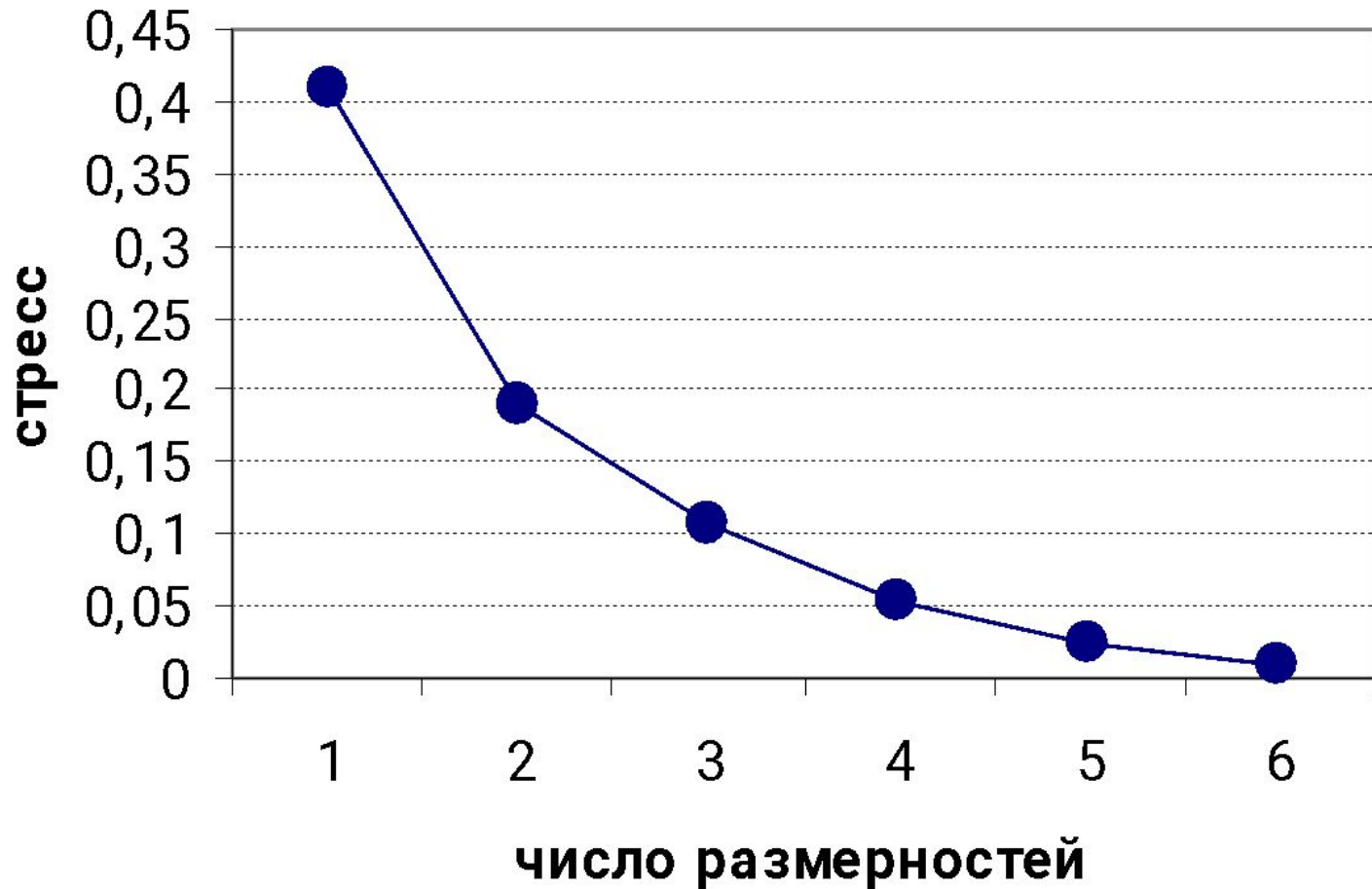
Для первого примера про города Беларуси:





Выбор размерности пространства

Для второго примера про страны:





Выбор размерности пространства

2) Вторым критерием для решения вопроса о размерности с целью интерпретации является "ясность" полученной конфигурации точек



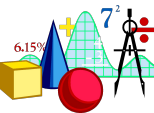


Преимущества

Часто информация, полученная от наблюдателя, носит неметрический характер, так как расстояния оцениваются по шкале порядка.



Итоговая же конфигурация воспроизводит метрические соотношения в расположении объектов!





Преимущества

Это связано с тем, что информация о различиях, содержащаяся в матрице субъективных оценок, хотя и является по сути порядковой, но обладает избыточностью, которая и позволяет восстановить метрические соотношения.





Преимущества

В общем случае, методы МНШ позволяют исследователю задать клиентам в анкете относительно ненавязчивые вопросы (например, "насколько похож товар фирмы А на товар фирмы В") и найти латентные переменные для этих анкет незаметно для респондентов



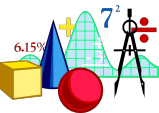


Полезная литература:

**К семинару по многомерному шкалированию
прочитать:**

**Лосик Г.В. «Исследование восприятия
гласных методом многомерного
шкалирования»**

ПЖ, том 13, № 2, 1992





Вот и все!

