

Programa de Pós-Graduação em Ecologia e Evolução – UFG

Métodos de Análise de Dados em Ecologia de Comunidades

Página do curso: www.ecologia.ufrgs.br/~adrimelo/div

Prof. Adriano Sanches Melo asm.adrimelo@gmail.com
Departamento de Ecologia
Universidade Federal de Goiás

Aula 4a

Índices de similaridade: Qualitativos: Presença e ausência

		amostra 1	
		+	-
amostra 2	+	a	b
	-	c	d

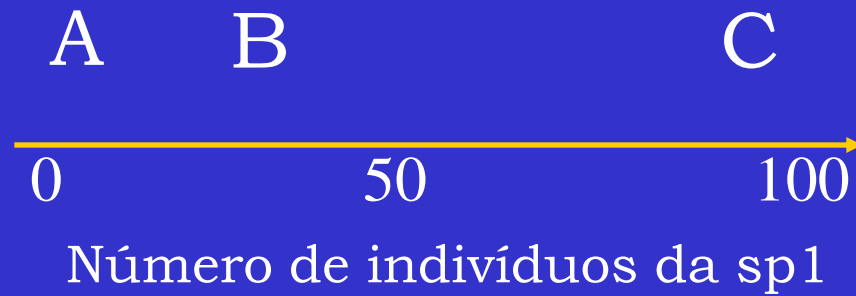
- Discordância quanto ao uso de “d”
- Geralmente variam entre 0 e 1

$$S_j = \frac{a}{a+b+c} \quad \text{Jaccard}$$

$$S_s = \frac{2a}{2a+b+c} \quad \text{Sorensen}$$

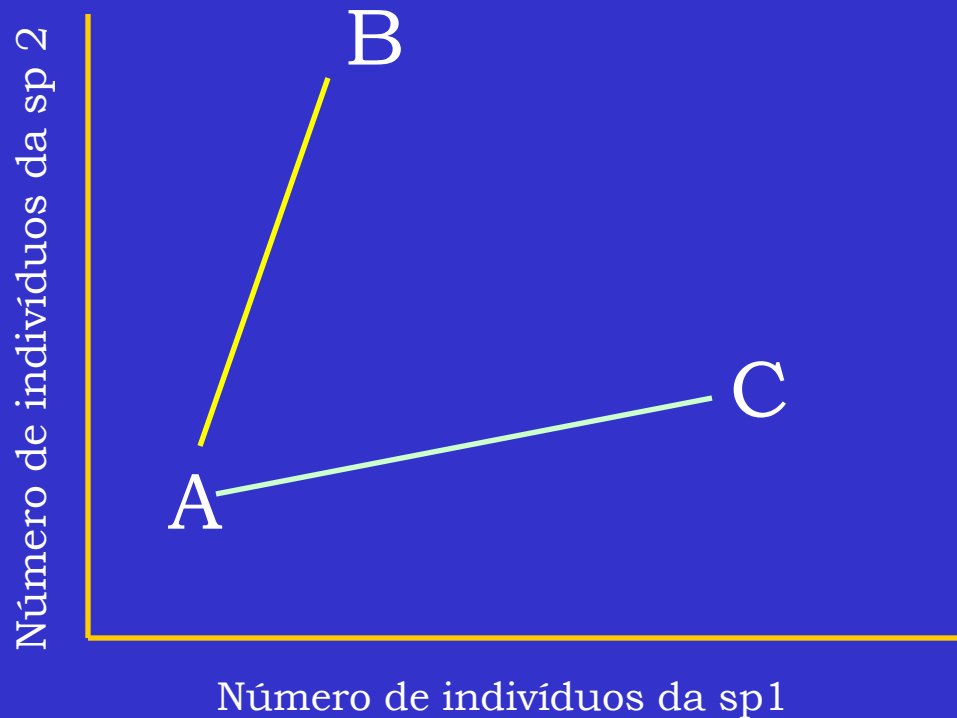
$$S_{sm} = \frac{a+d}{a+b+c+d} \quad \text{Concordância simples (“simple matching”)}$$

Índices de similaridade: Quantitativos: também abundância



Índices de similaridade: Quantitativos: também abundância

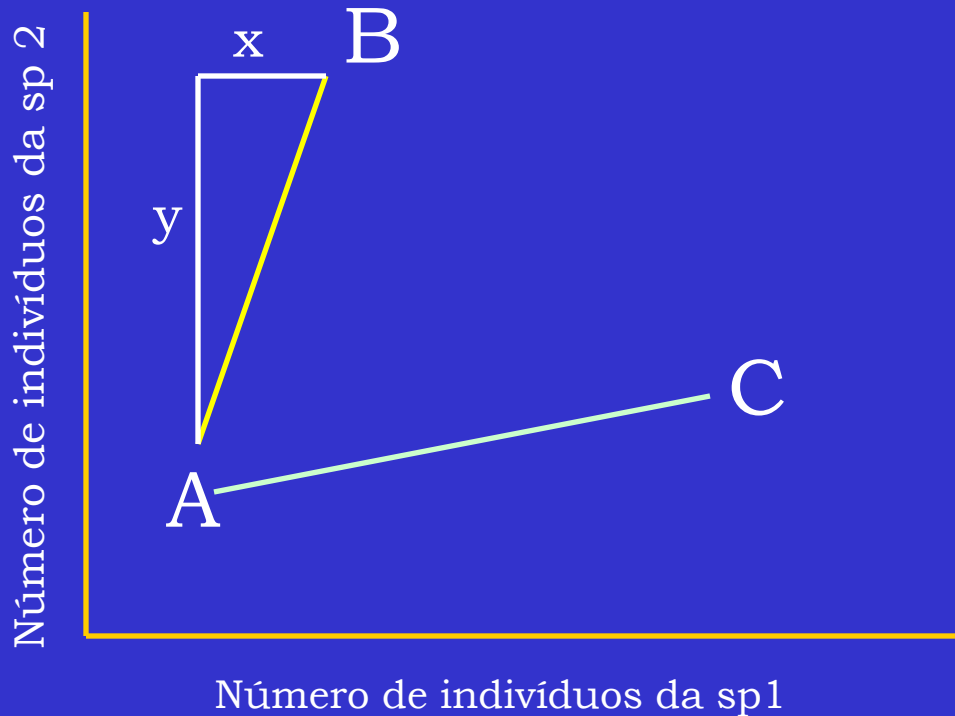
Distância Euclidiana



Índices de similaridade: Quantitativos: também abundância

Distância Euclidiana

$$D_{jk} = \sqrt{x^2 + y^2}$$

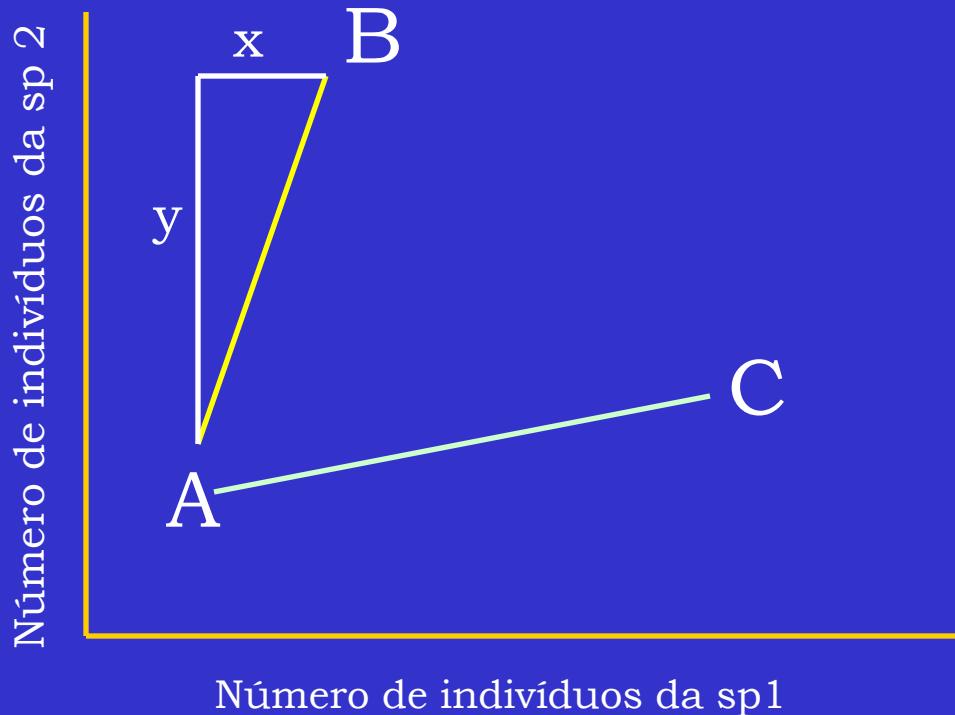


Índices de similaridade: Quantitativos: também abundância

Distância Euclidiana

$$D_{jk} = \sqrt{x^2 + y^2}$$

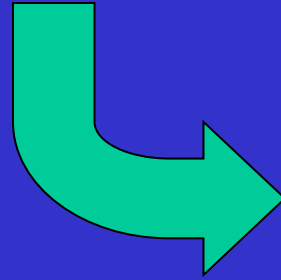
E para 3, 4, 5 etc espécies?



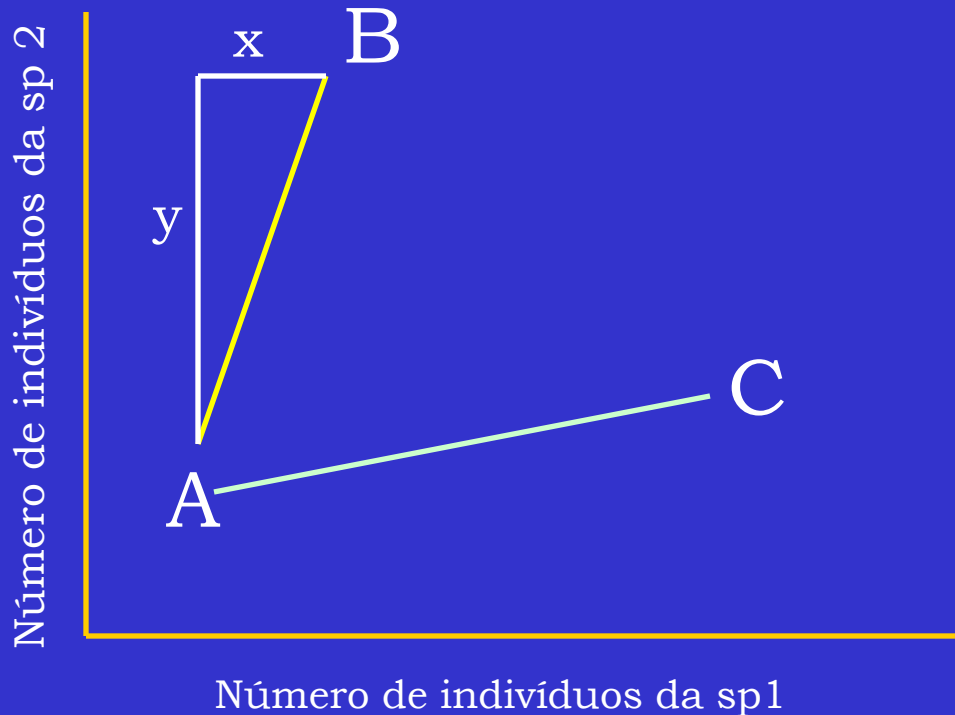
Índices de similaridade: Quantitativos: também abundância

Distância Euclidiana

$$D_{jk} = \sqrt{x^2 + y^2}$$



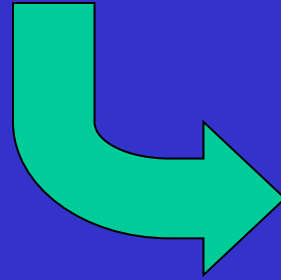
$$\Delta_{jk} = \sqrt{\sum_{i=1}^s (X_{ij} - X_{ik})^2}$$



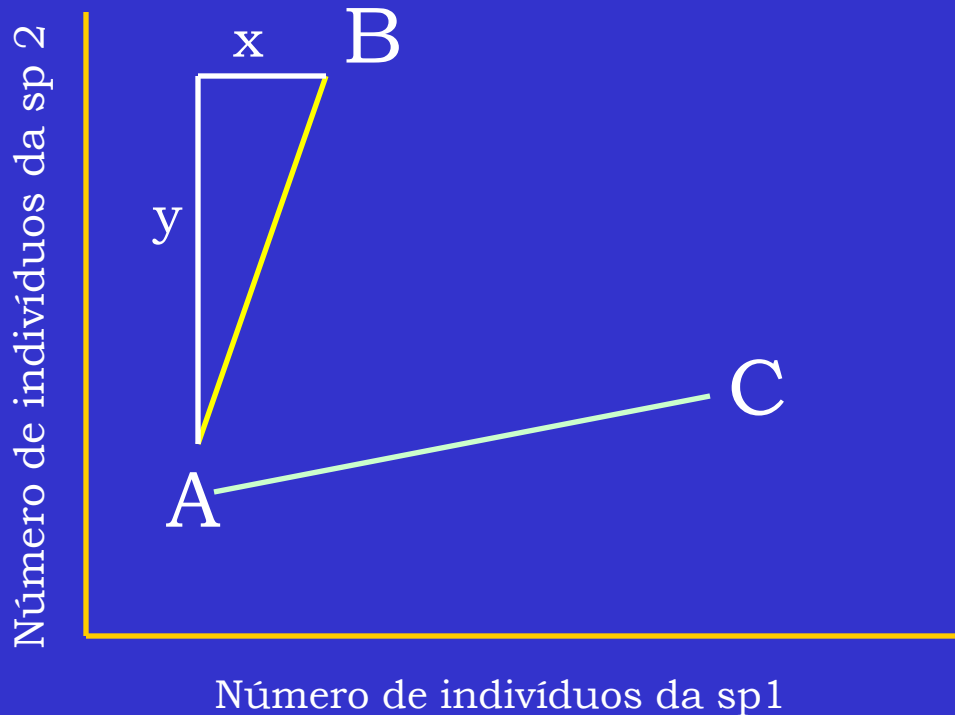
Índices de similaridade: Quantitativos: também abundância

Distância Euclidiana

$$D_{jk} = \sqrt{x^2 + y^2}$$



$$D_{jk} = \sqrt{\sum_{i=1}^s (X_{ij} - X_{ik})^2}$$



varia de 0 ao ∞

Índices de similaridade: Quantitativos: também abundância

Distância Euclidiana

$$\Delta_{jk} = \sqrt{\sum_{i=1}^s (X_{ij} - X_{ik})^2}$$

varia de 0 ao ∞

Distância Manhattan
("city block")

$$\Delta_{jk} = \sum_{i=1}^s |X_{ij} - X_{ik}|$$



Distância Bray-Curtis

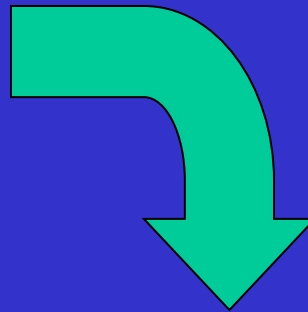
$$d_m(j, k) = \frac{\sum_{i=1}^s |X_{ij} - X_{ik}|}{\sum_{i=1}^s (X_{ij} + X_{ik})}$$

varia de 0 a 1

Índices de similaridade

Dados originais

	sp 1	sp 2	sp 3	sp 4	sp 5
a1	5	2	5	2	1
a2	0	1	3	2	1
a3	2	1	3	2	1
b1	5	20	6	5	5
b2	12	19	4	7	11
b3	11	21	5	7	10



Matriz de dissimilaridade
(Bray-Curtis)

	a1	a2	a3	b1	b2	b3
a1	0					
a2	0,364	0				
a3	0,250	0,125	0			
b1	0,464	0,708	0,640	0		
b2	0,588	0,766	0,709	0,191	0	
b3	0,565	0,770	0,714	0,157	0,047	0

Teste de Mantel - Correlação entre matrizes de distância

Existe relação entre dissimilaridade em composição de espécies e distância geográfica?

Locais distintos em condições ambientais possuem floras distintas?

Qual distância explica melhor dissimilaridades biológicas:

- a) distância geográfica ou pelo corredor de riachos?
- b) distância geográfica ou por conectividade entre lagos?

Dissimilaridade biológica

	a1	a2	a3	b1	b2	b3
a1	0					
a2	0,364	0				
a3	0,250	0,125	0			
b1	0,464	0,708	0,640	0		
b2	0,588	0,766	0,709	0,191	0	
b3	0,565	0,770	0,714	0,157	0,047	0

Distância geográfica

	a1	a2	A3	b1	b2	b3
a1	0					
a2	3,64	0				
a3	3,25	1,25	0			
b1	3,46	9,08	6,99	0		
b2	7,08	6,66	6,00	0,98	0	
b3	3,65	8,98	7,55	1,01	0,07	0

Teste de Mantel - Correlação entre matrizes de distância

Dissimilaridade biológica

	a1	a2	a3	b1	b2	b3
a1	0					
a2	0,364	0				
a3	0,250	0,125	0			
b1	0,464	0,708	0,640	0		
b2	0,588	0,766	0,709	0,191	0	
b3	0,565	0,770	0,714	0,157	0,047	0

Distância geográfica

	a1	a2	A3	b1	b2	b3
a1	0					
a2	3,64	0				
a3	3,25	1,25	0			
b1	3,46	9,08	6,99	0		
b2	7,08	6,66	6,00	0,98	0	
b3	3,65	8,98	7,55	1,01	0,07	0

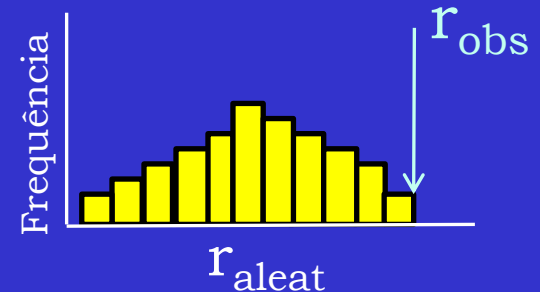


0,364 3,64
0,250 3,25
0,464 3,46
0,588 7,08
0,565 3,65
0,125 1,25
...
0,047 0,07

Correlação
Teste de aleatorização

Testes de Aleatorização – Mantel e todos outros!

- 1) Escolha uma estatística que mensure padrão que se quer testar
→ correlação de Pearson
- 2) Calcule estatística para conjunto de dados observados → r_{obs}
- 3) Aleatorize dados → mude posição de valores dentro de uma das colunas
- 4) Calcule estatística para conjunto de dados aleatorizados → r_{aleat}
- 5) Repita passos 3 e 4 muitas vezes → 1000 vezes
- 6) Compare estatística observada com estatísticas obtidas com aleatorização

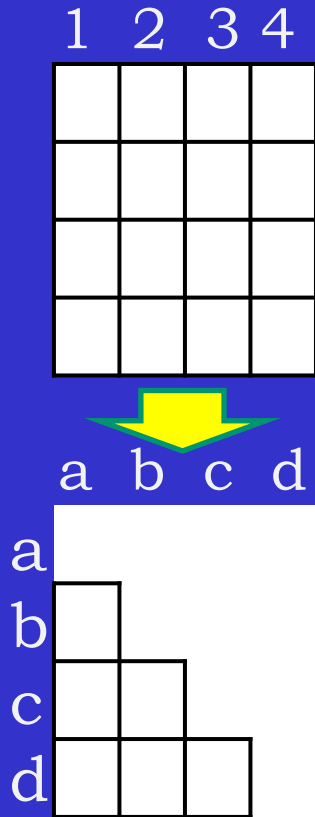


- 7) Obtenha probabilidade de estatística observada ser semelhante a estatísticas obtidas com dados aleatorizados

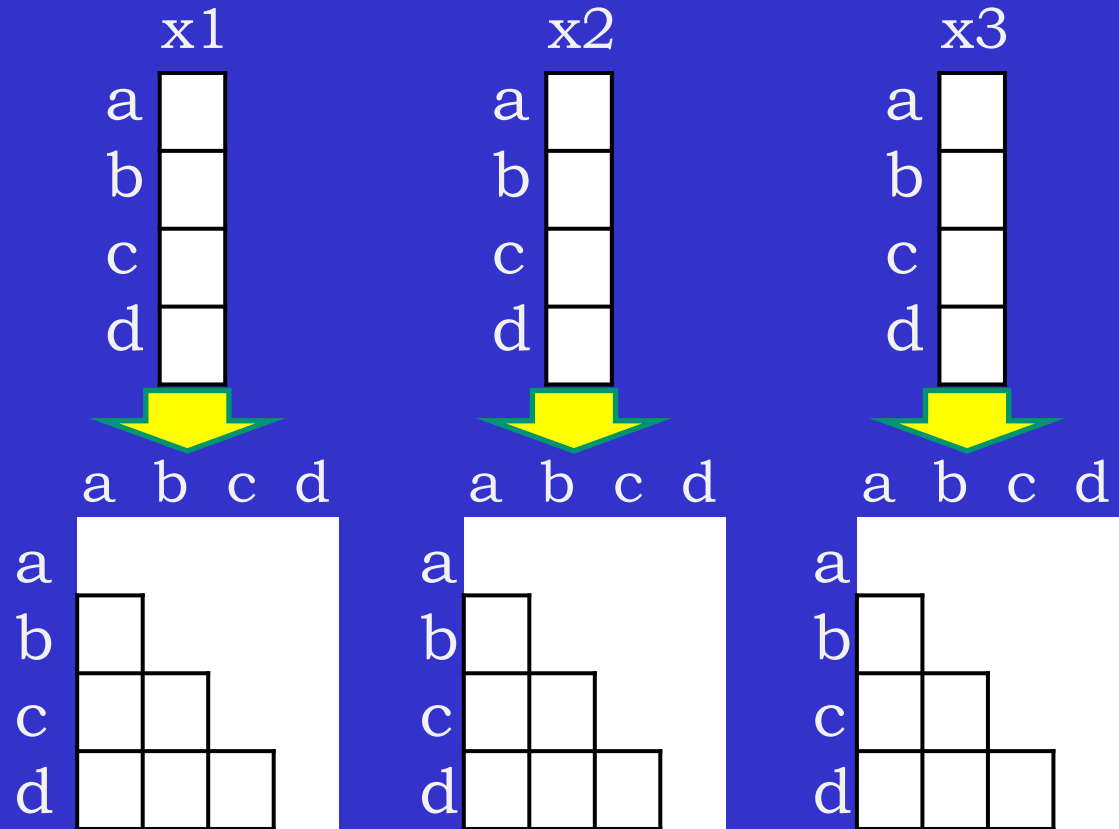
No nosso caso temos teste monocaudal e portanto

$$p = \frac{(\text{Frequência de } r_{aleat} \geq r_{obs}) + 1}{(\text{Número aleatorizações} + 1)}$$

Biológica



Ambientais



Com qual matriz ambiental existe maior correlação?

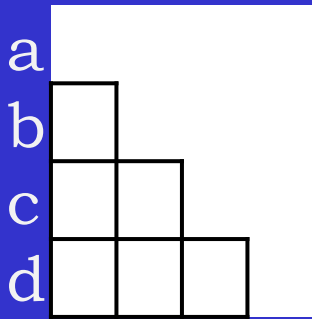
Análise Bioenv

Biológica

1 2 3 4



a b c d



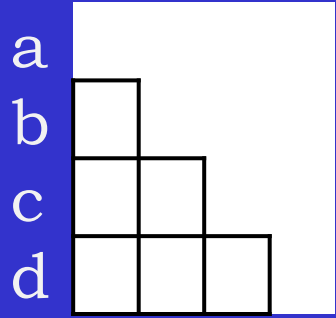
Ambientais

x1 x2

a		
b		
c		
d		



a b c d

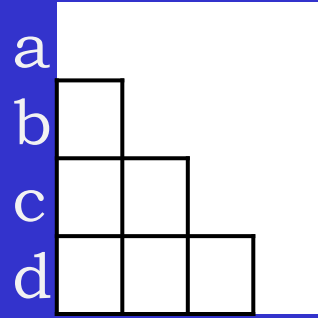


x1 x3

a		
b		
c		
d		



a b c d

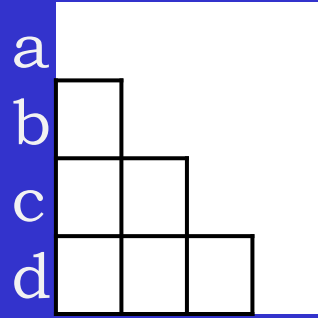


x2 x3

a		
b		
c		
d		



a b c d



Com qual matriz ambiental existe maior correlação?

Análise Bioenv

Biológica

1 2 3 4



a b c d

a				
b				
c				
d				

Ambientais

x1 x2 x3

a			
b			
c			
d			



a b c d

a				
b				
c				
d				

Qual correlação?

Análise Bioenv - Exemplo de resultado

Tam.Modelo	Modelo	Correlação
1	cond	0,457
2	cond + links	0,589
3	cond + links + order	0,619
4	cond + links + order + width	0,603
5	cond + links + order + width + flow	0,596
6	cond + links + order + width + flow + alt	0,566
7	cond + links + order + width + flow + alt + cover	0,515