Adatelemzés és mérések

Budapest University of Technology and Economics
Fault Tolerant Systems Research Group

BENCHMARKING

Why benchmark?

Benchmarking

- Célok: szoftver/hardver eszközök teljesítményének összehasonlítása
  - Döntéstámogatás
    - melyiket előnyösebb megvenni/telepíteni
    - mekkora terhelésre elég a meglévő rendszer
  - Teljesítménytesztelés
    - kell-e még teljesítményen javítani és hol (fejlesztésnél)
    - optimális-e egy konkrét beállítás
    - van-e egy beállításnak teljesítményre gyakorolt hatása

Elvárások

- Ismételhetőség
  - Repeatability
- Reprodukálhatóság
  - Reproducilibity
- Relevancia
- Szabványok/megállapodások betartása
- Általánosított felhasználói eset
  - Átlag felhasználó számára értelmezhető legyen az eredmény

Benchmark terhelési modellek

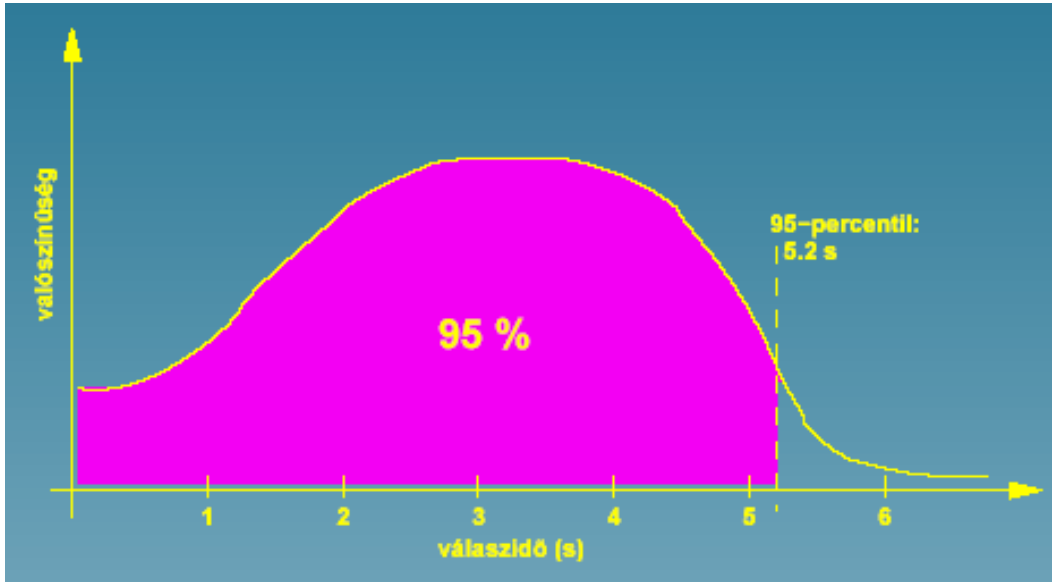
- Tudományos/műszaki rendszerek
  - nagy mennyiségű adat feldolgozása (number crunching)
  - párhuzamos módszerek
- Tranzakciókezelés (OLTP)
  - kliens-szerver környezet
  - sok gyors, párhuzamos tranzakció
- Batch jellegű adatfeldolgozás
  - riport készítés nagy mennyiségű adatból
- Döntéstámogatás
  - kevés, bonyolult lekérdezés
  - ad hoc műveletek
  - sok adat (pl. OLAP)
- Virtualizáció

Mérendő paraméterek (Metrikák)

- Futási idő
  - kezdet, vég?
  - eloszlás
  - CPU, I/O, hálózat,…
- Tranzakiósebesség
  - rendszer reakcióideje
  - akár egymásba ágyazott tranzakciók
- Áteresztőképesség
  - ***feldolgozott*** adatmennyiség / futási idő
  - terhelés függvényében

Mérendő paraméterek (2)

- Válaszidő
  - terhelés függvényében
    - felhasználók
    - tranzakciók száma, stb.
- X-Percentil
  - Egy adott halmaz X százaléka ez alatt az érték alatt van



Benchmark elvégzése

- Relevancia biztosítása
  - Tényleg azt az alkalmazást mérjük, amit kell
  - Terhelésgenerálás jellege közelítse a valódi terhelést
  - Minimalizáljuk a zavaró tényezőket

SPEC benchmarkok

- http://www.spec.org/benchmarks.html
  - Standard Performance Evaluation Corp.
- Erőforrás és alkalmazás szintű benchmarkok
  - CPU
  - Alkalmazások
  - Levelező szerverek
  - Web szerverek stb.
- Benchmark: megrendelhető szolgáltatás

SPEC CPU2006

- CPU intenzív
- CINT2006
  - Számításigényes egész számos
- CFP2006
  - Lebegőpontos

- Példa:

cpu2006.html

CINT2006 és CFP2006 terhelésgenerátorok

- CINT2006 :

400.perlbench	C	Programming Language
401.bzip2	C	Compression
403.gcc	C	C Compiler
429.mcf	C	Combinatorial Optimization
445.gobmk	C	Artificial Intelligence
456.hmmer	C	Search Gene Sequence
458.sjeng	C	Artificial Intelligence
462.libquantum	C	Physics / Quantum Computing
464.h264ref	C	Video Compression
471.omnetpp	C++	Discrete Event Simulation
473.astar	C++	Path-finding Algorithms
483.xalancbmk	C++	XML Processing

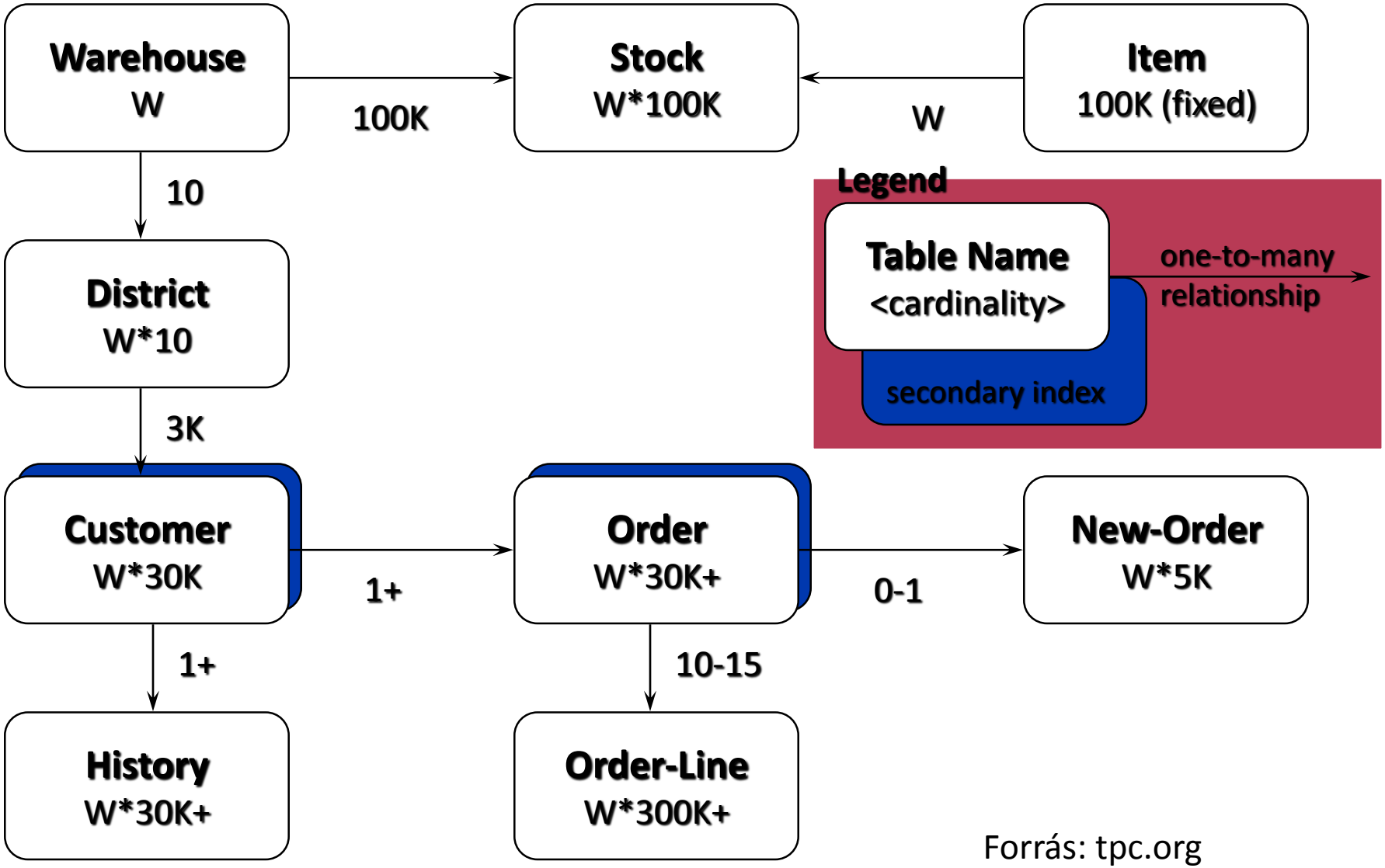
- CFP2006:

SPEC CFP2006 Benchmark Descriptions.htm

A TPC benchmark

- Adatbáziskezelő rendszerek mérése
  - RDBMS+OS+HW
- Mérési környezet
  - Mintaadatbázis: Ügyfelek és megrendelések
  - 5 fajta tranzakció (lekérdezés/módosítás) vegyesen
  - Felső korlát a futási időre
  - Valós körülmények: ACID tranzakciók, felhasználói gondolkodási idők
- Mért adatok
  - Áteresztőképesség (tpmC)
  - „Hatékonyság” ($/tpmC)

TPC-C séma



Forrás: tpc.org

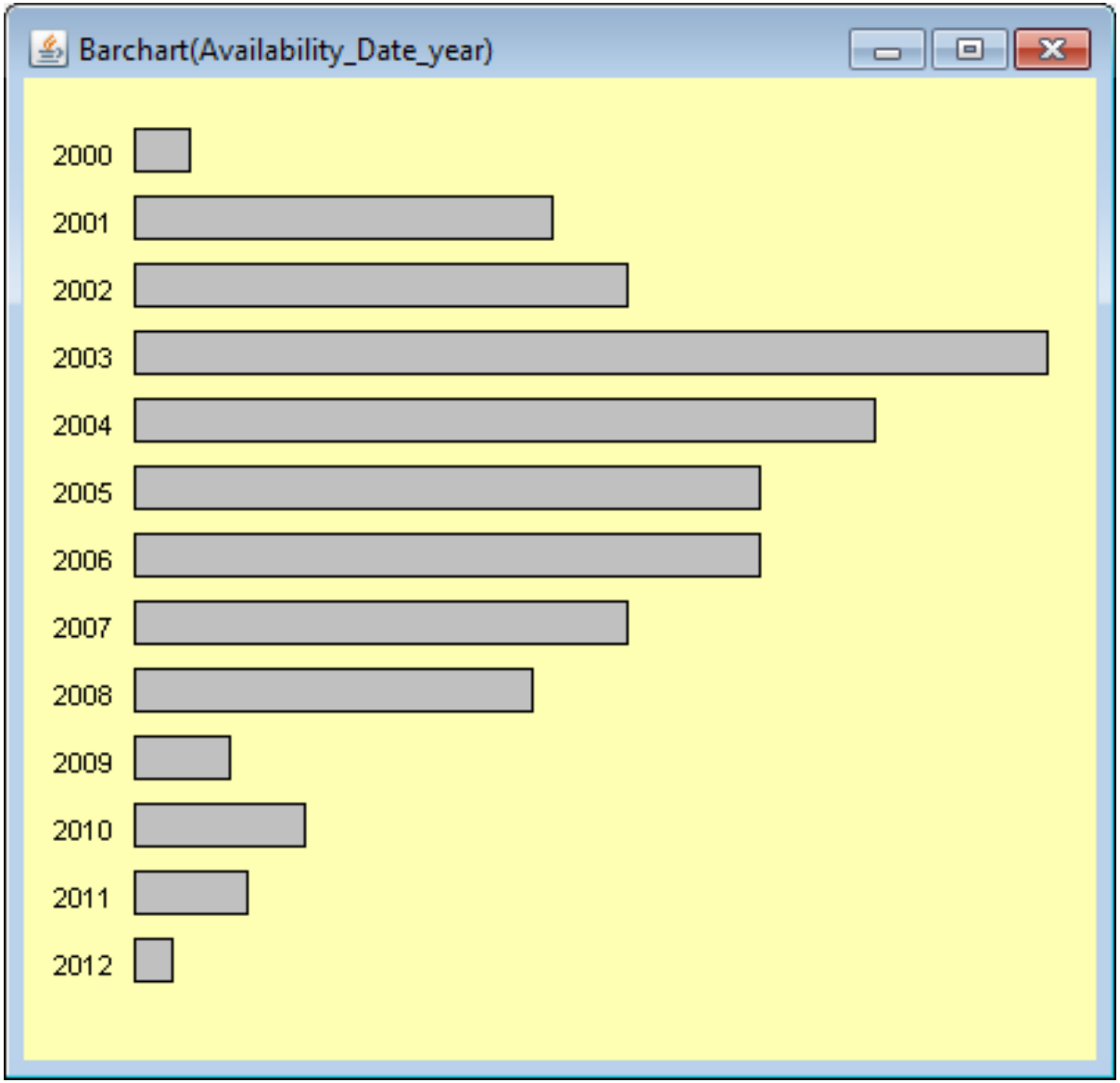
Mielőtt elemeznénk: tisztítunk

- A kiinduló adathalmazunk:

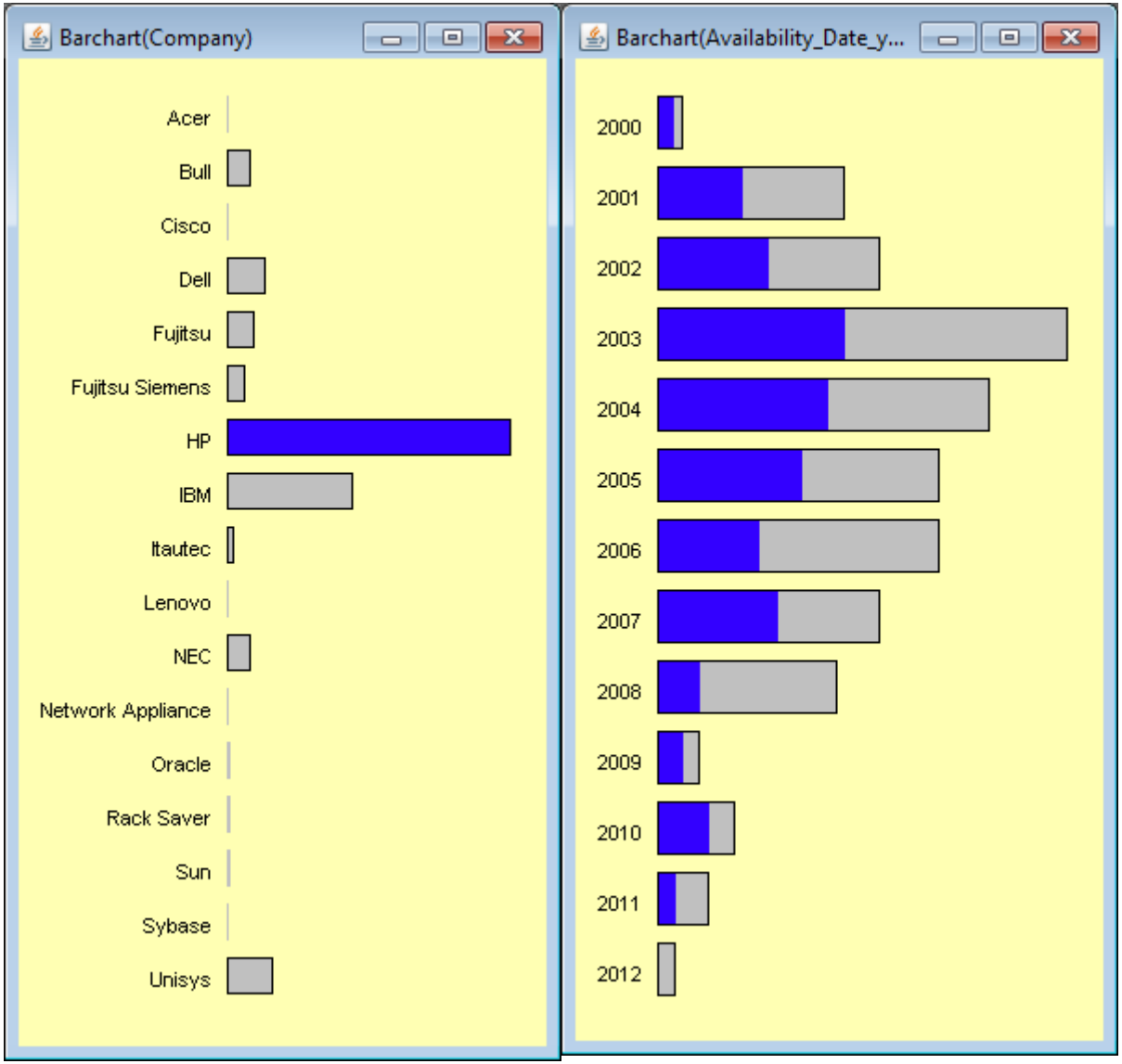
	A	B	C	D	E	F	G	H	I	J	K
1	TPC-C BENCHMARK RESULTS										
2	These results are valid as of date 6/12/2012 10:04:24 PM										
3											
4		TPC-C Results - Revision 5.X									
5											
6	Company	System	Spec. Revision	tpmC	Price/Perf	Total Sys. Cost	Currency	Database Software	Operating System	TP Monitor	Server CPU Type
7	Acer	Altos R710	5.5	66543	12.42	826507.55	AUD	Microsoft SQL Server	Microsoft Windows Serv	Microsoft CO	Intel Xeon - 3.6 GHz
8	Bull	Bull Escal	5.9	6085166	2.81	17127928	USD	IBM DB2 9.5	IBM AIX 5L V5.3	Microsoft CO	IBM POWER6 - 5.0
9	Bull	Bull Escal	5.9	629159	2.49	1566664	USD	IBM DB2 9.5 Enterpri	IBM AIX 5L V5.3	Microsoft CO	IBM POWER6 - 4.2
10	Bull	Bull Escal	5.8	1616162	3.54	5716286	USD	IBM DB2 9.1	IBM AIX 5L V5.3	Microsoft CO	IBM POWER6 - 4.7
11	Bull	Bull Escal	5.8	404462	3.51	1417121	USD	Oracle Database 10g	IBM AIX 5L V5.3	Microsoft CO	IBM POWER6 - 4.7

- Felesleges adatok:
  - Sorok (pl. a kezdő sorok, és az állomány végén lévő sorok, amelyek nem kapcsolódnak az eredményekhez)
  - Oszlopok (pl. Server CPU Type nekünk most nem kell)
  - Eltérő valutákban megadott költségek

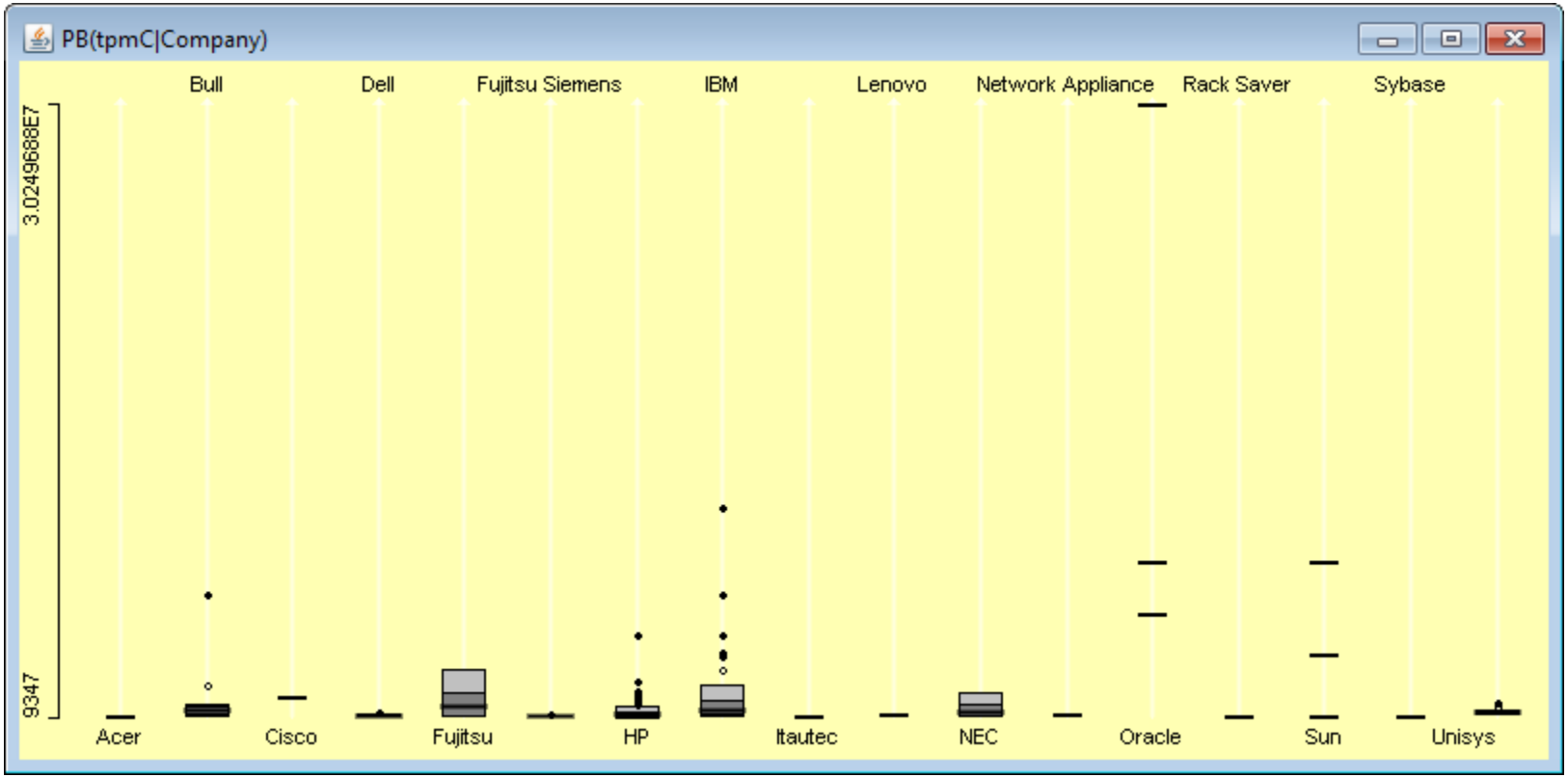
A benchmark mely évek eredményeit tartalmazza?



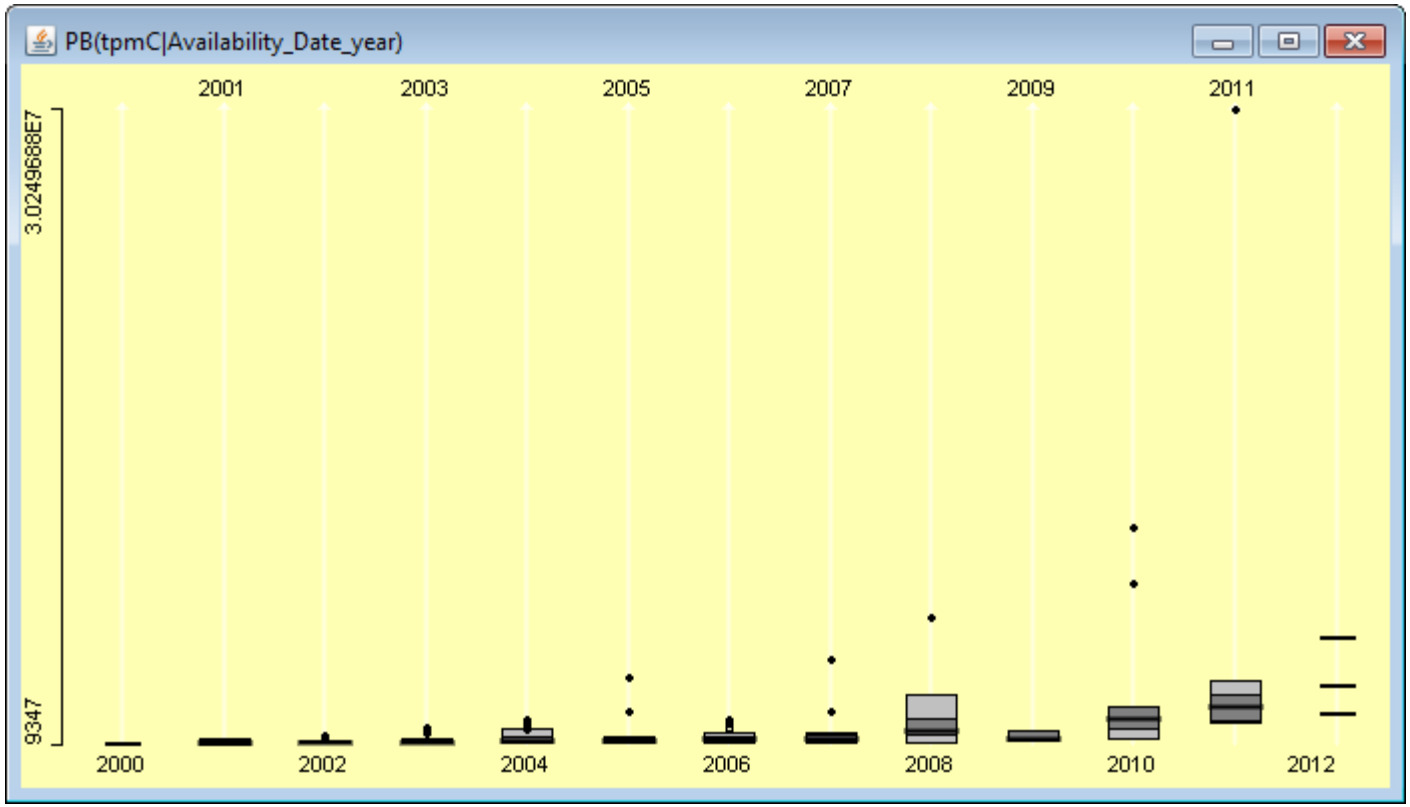
Mikor voltak aktívak a beszállítók?



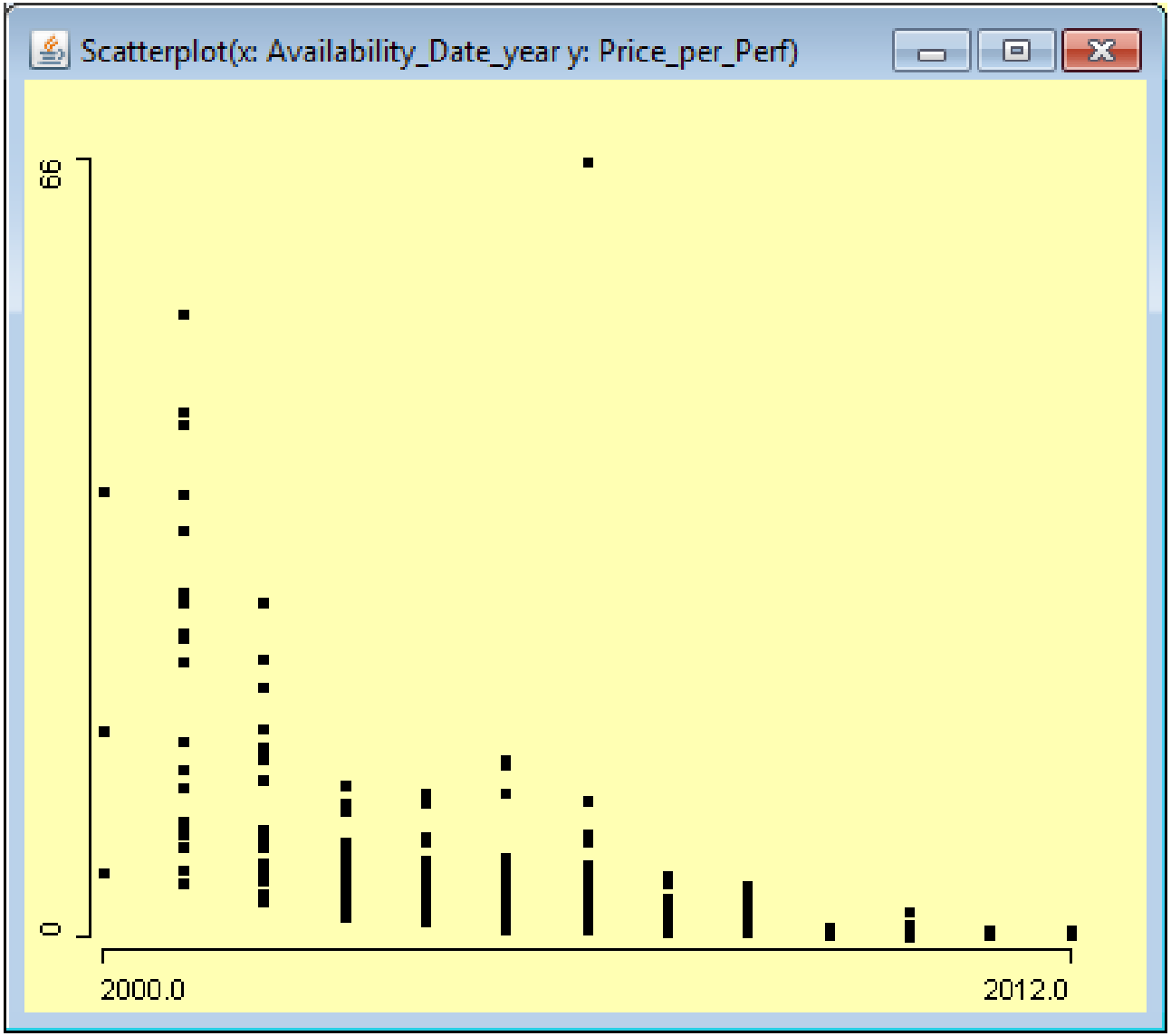
Hogyan alakulnak a teljesítmény mérőszámok?



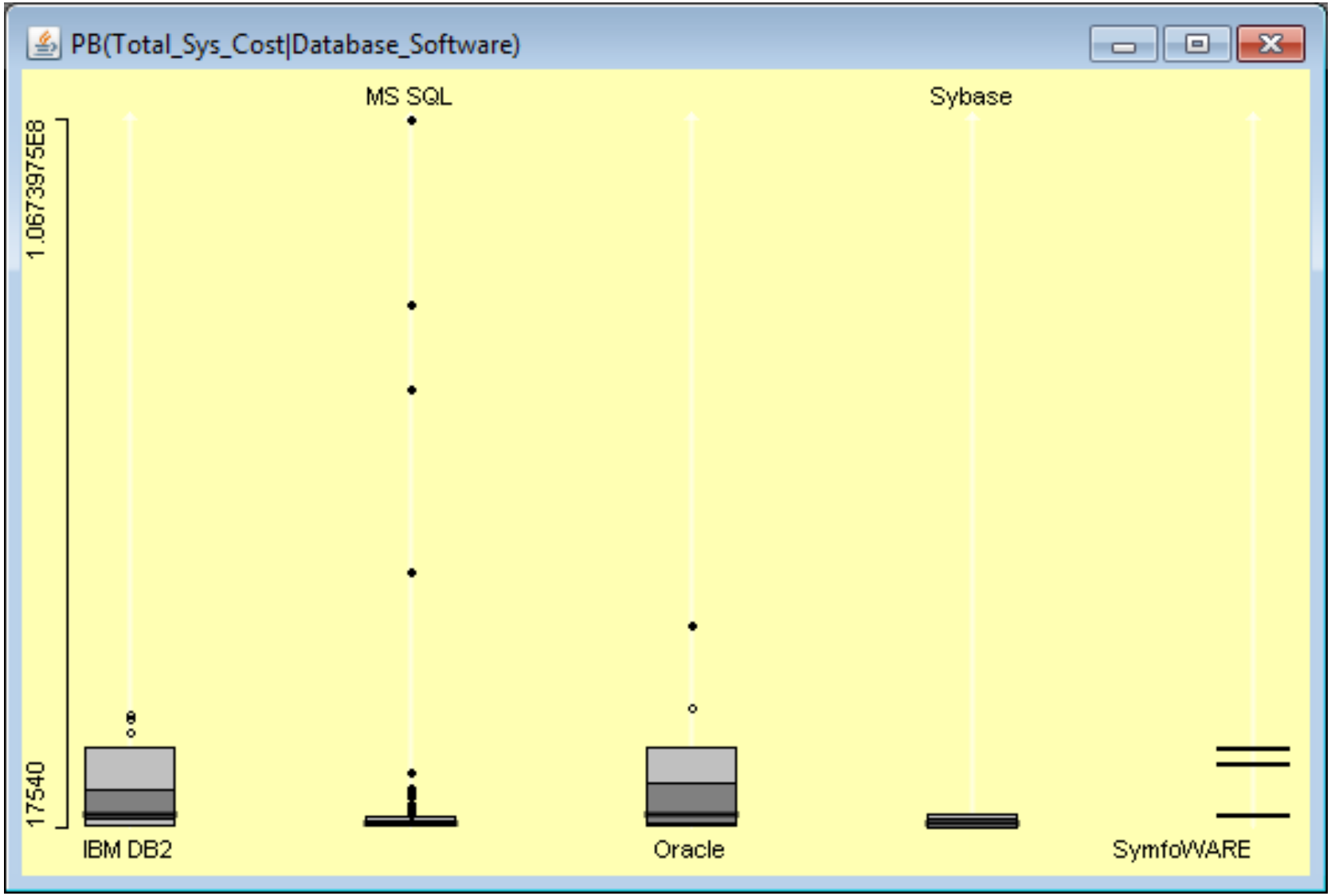
Időben a teljesítmény hogyan változott?



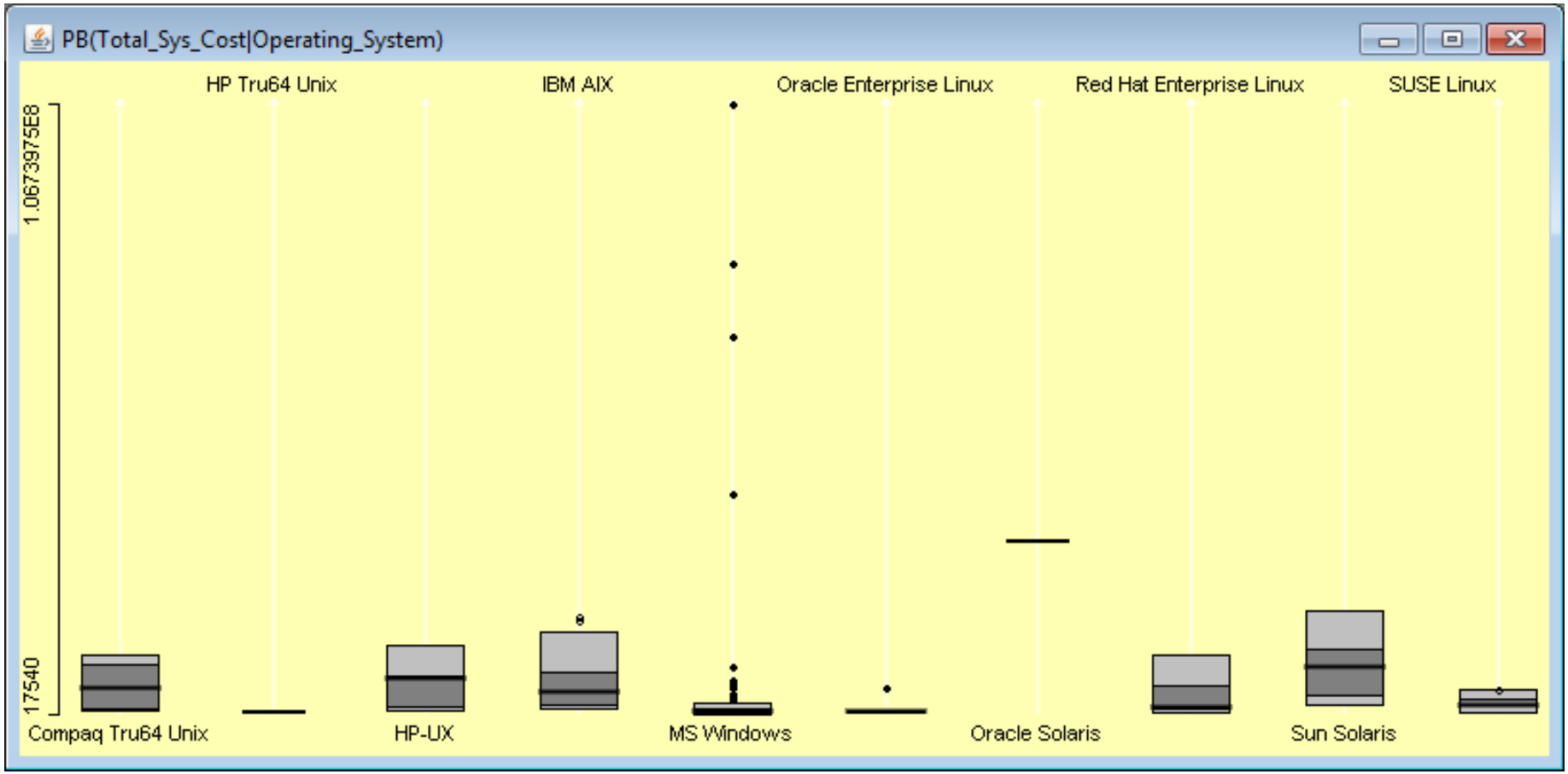
Hogyan alakultak mindeközben az árak?



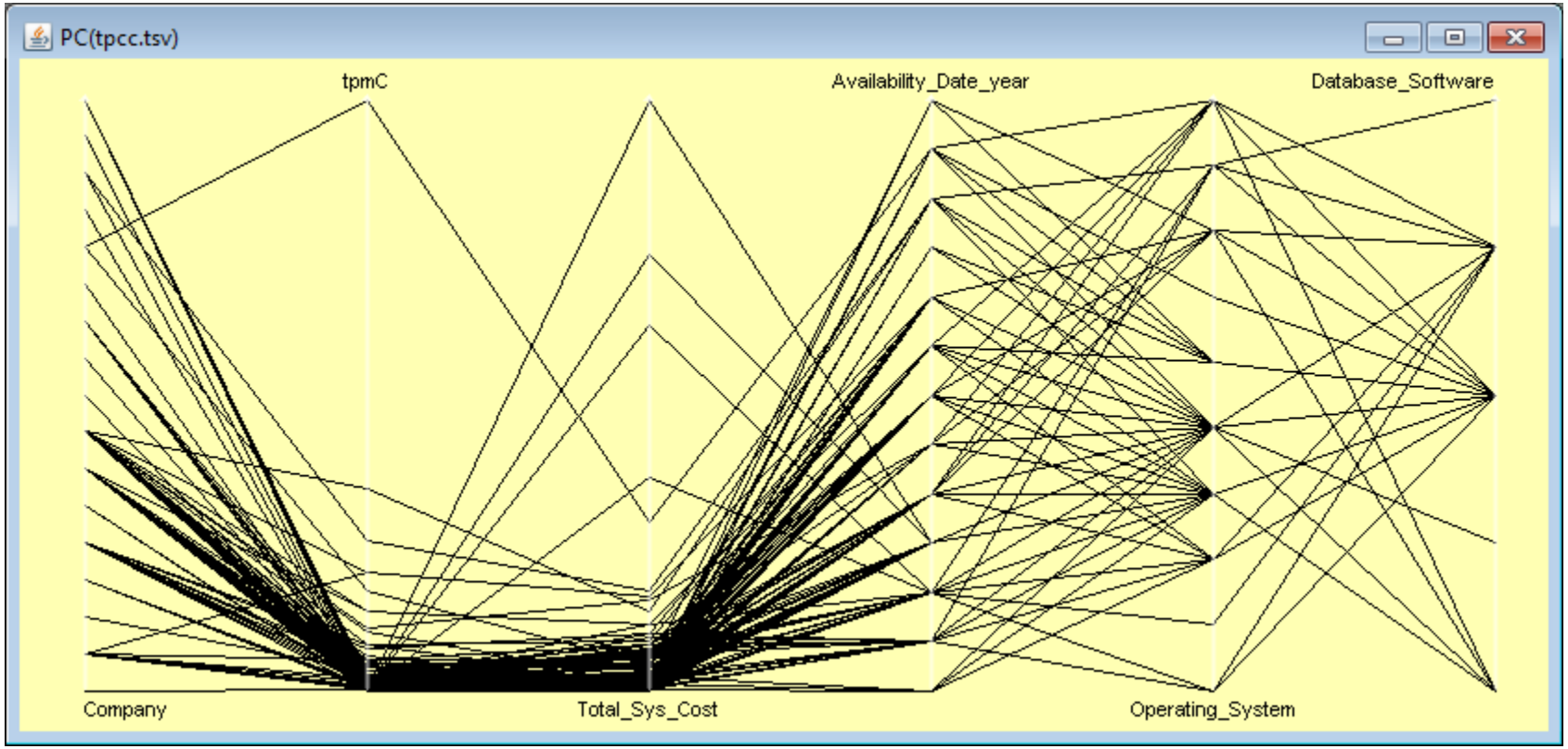
Milyen adatbázis-kezelő SW-t válasszunk?



Milyen OS-t válasszunk?



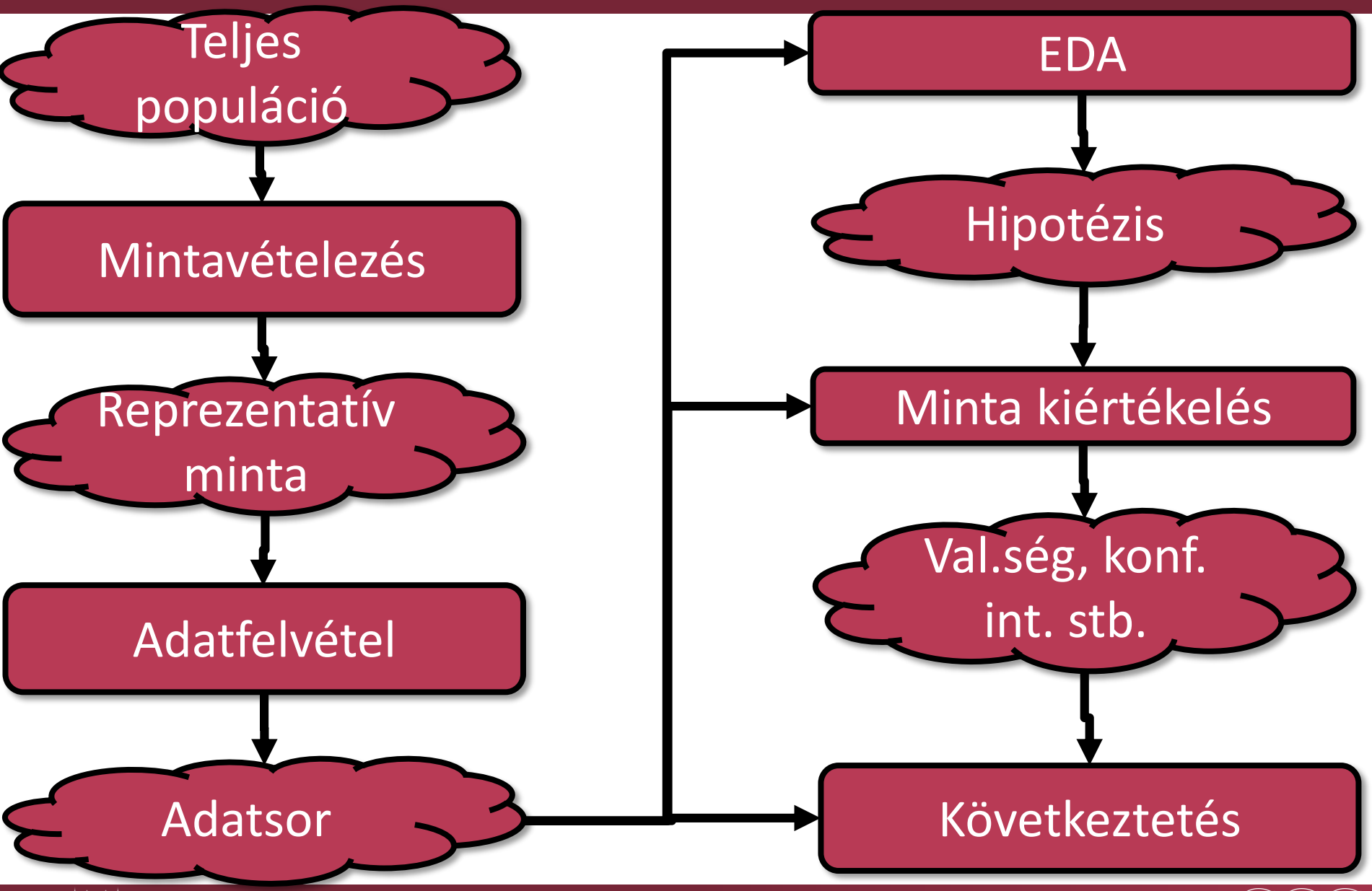
A „big picture”



KÖVETKEZTETŐ STATISZTIKA

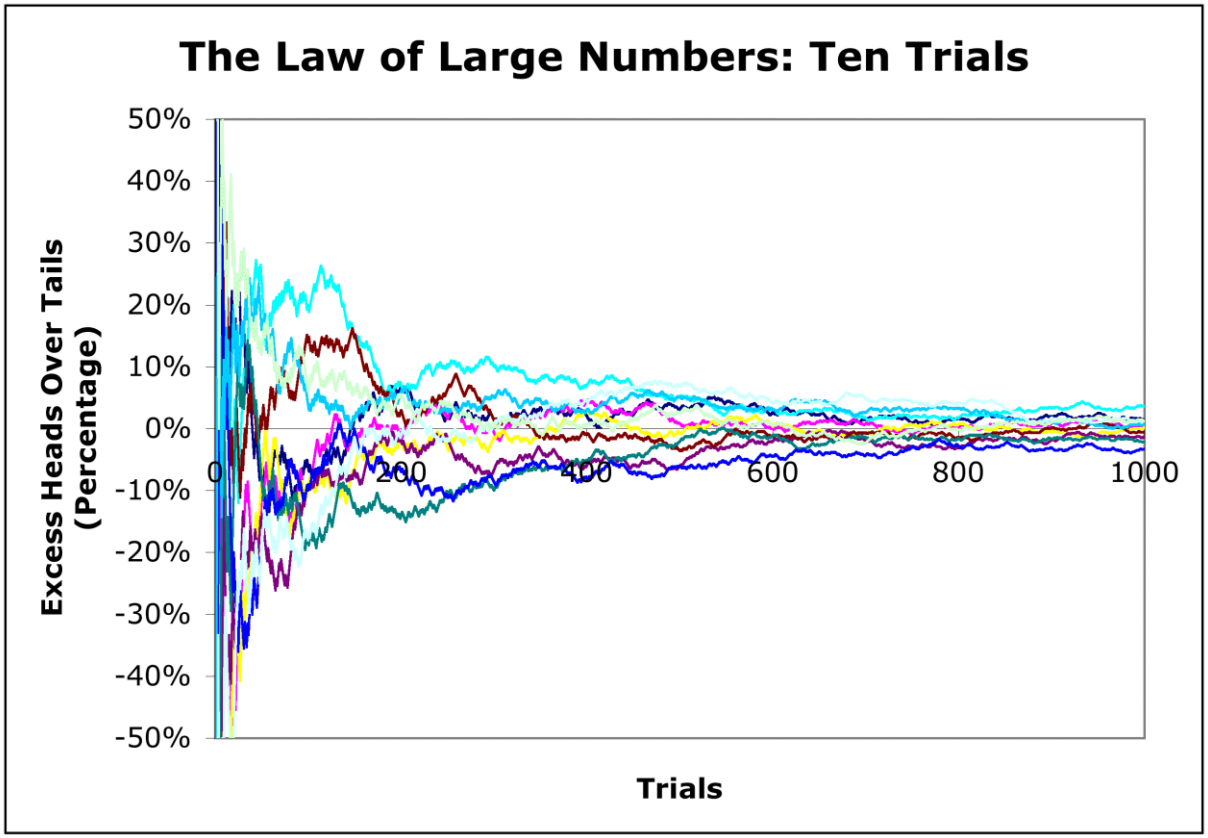
Következtető statisztika

Következtető statisztika



Ökölszabályok

- LLN (Law of Large Numbers)
  - Ha a kísérletek száma tart a végtelenhez, az előfordulási gyakoriság az elméleti valószínűséghez konvergál



?

Magyarországi kamaszlányok

Békés

Heves

Vas

$\bar{x}_{Békés}$

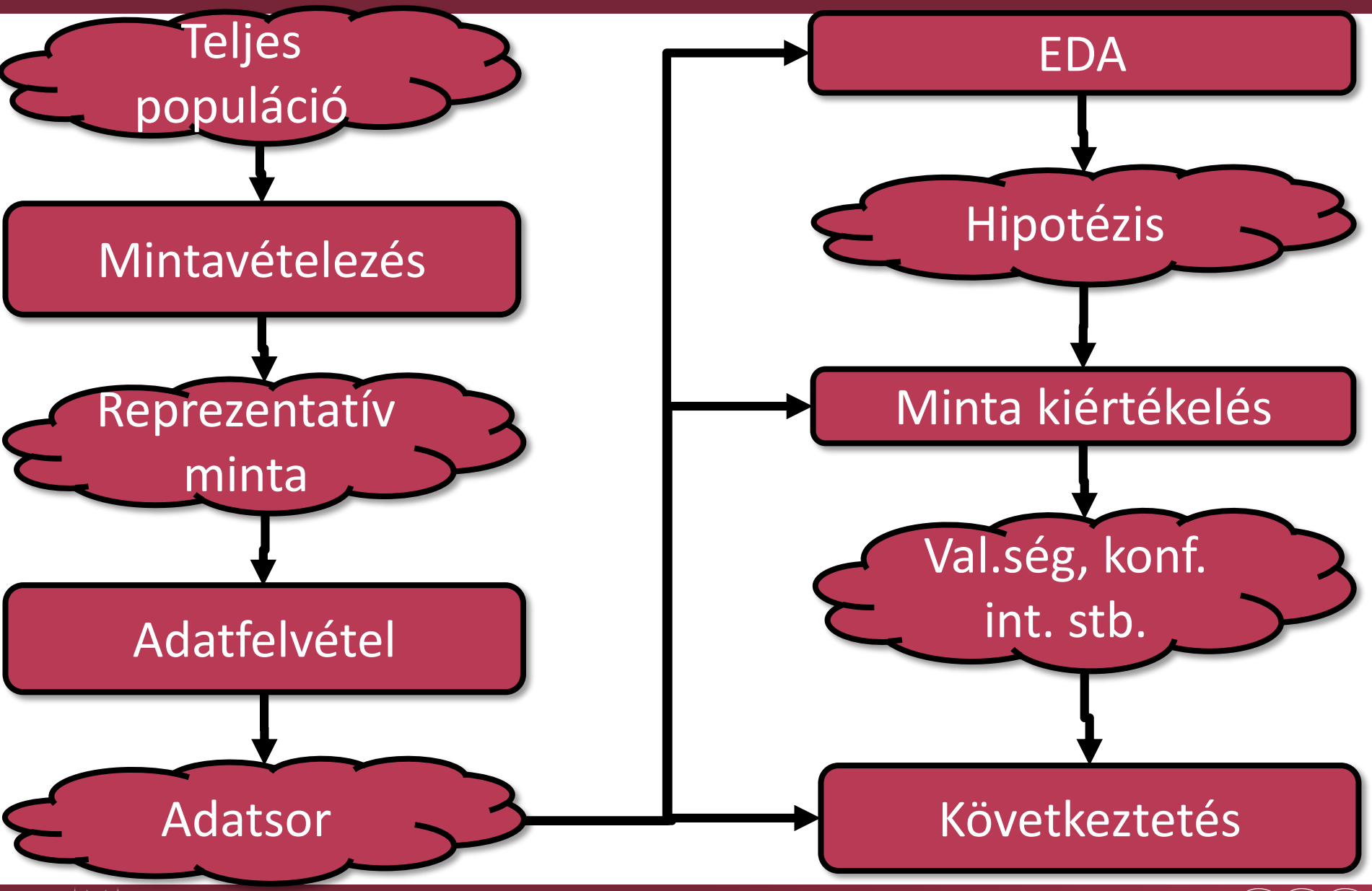
$\bar{x}_{Heves}$

$\bar{x}_{Vas}$

$$\mu = \frac{X_1 + X_2 + \ \dots + X_N}{N}$$

$$\mu \approx mean(\bar{x})$$

Következtető statisztika



Következtetés

- Döntési bemenet
  - Valami küszöbérték
- Adatsor típusa
  - Megfigyelési tanulmány (observational study)
  - Kísérlet (experiment)

Különbség: a *köztes változók* eliminálása

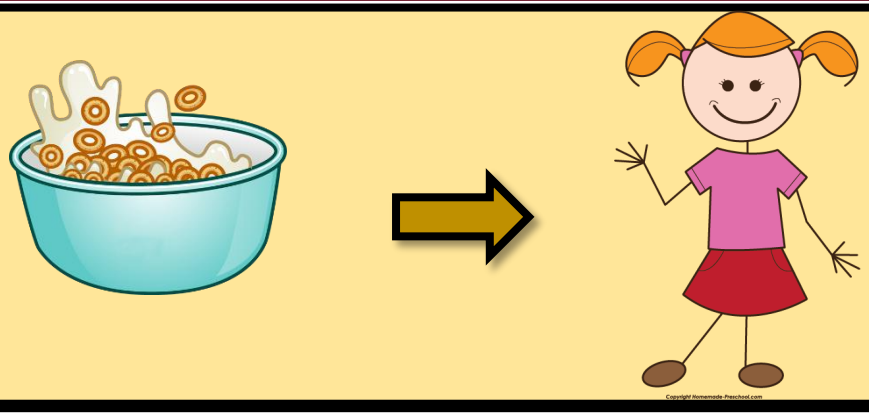
Esettanulmány

„Girls who ate breakfast of any type had a lower average body mass index, a common obesity gauge, than those who said they didn't. The index was even lower for girls who said they ate cereal for breakfast.„

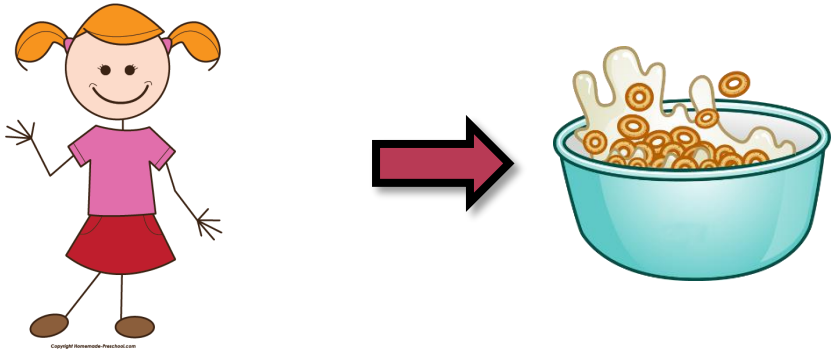
Forrás: http://usatoday30.usatoday.com/news/health/2005-09-08-cereal-slimming_x.htm

Esettanulmány

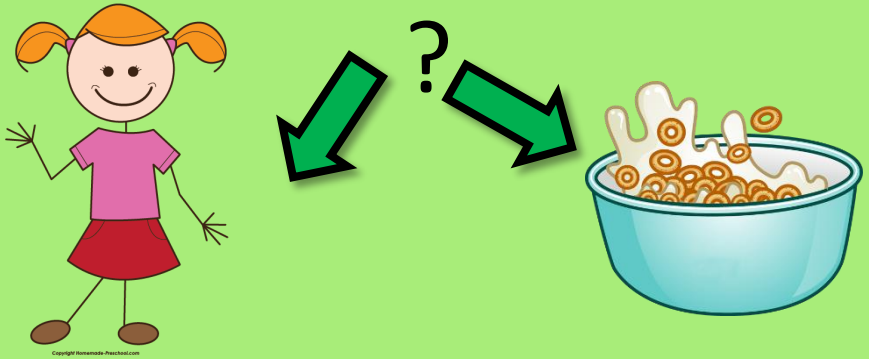
1. „Breakfast, cereal keep girls slim”



2. „Being slim causes girls to eat breakfast„



3. „A confounding variable is responsible for both”



Következtetés

- Döntési bemenet
  - Valami küszöbérték
- Adatsor típusa
  - Megfigyelési tanulmány (observational study)
    - A köztes változók kiléte bizonytalan
    - Csak korreláció, kauzális következtetések nem
  - Kísérlet (experiment)
    - A köztes változókat kiszűrtük (mintavételezés!)
    - Kauzális következtetések is