Modellek paraméterezése: regresszió, benchmarkok

Rendszermodellezés

Budapesti Műszaki és Gazdaságtudományi Egyetem
Méréstechnika és Információs Rendszerek Tanszék

MŰEGYETEM 1782

Teljesítményanalízis megközelítések

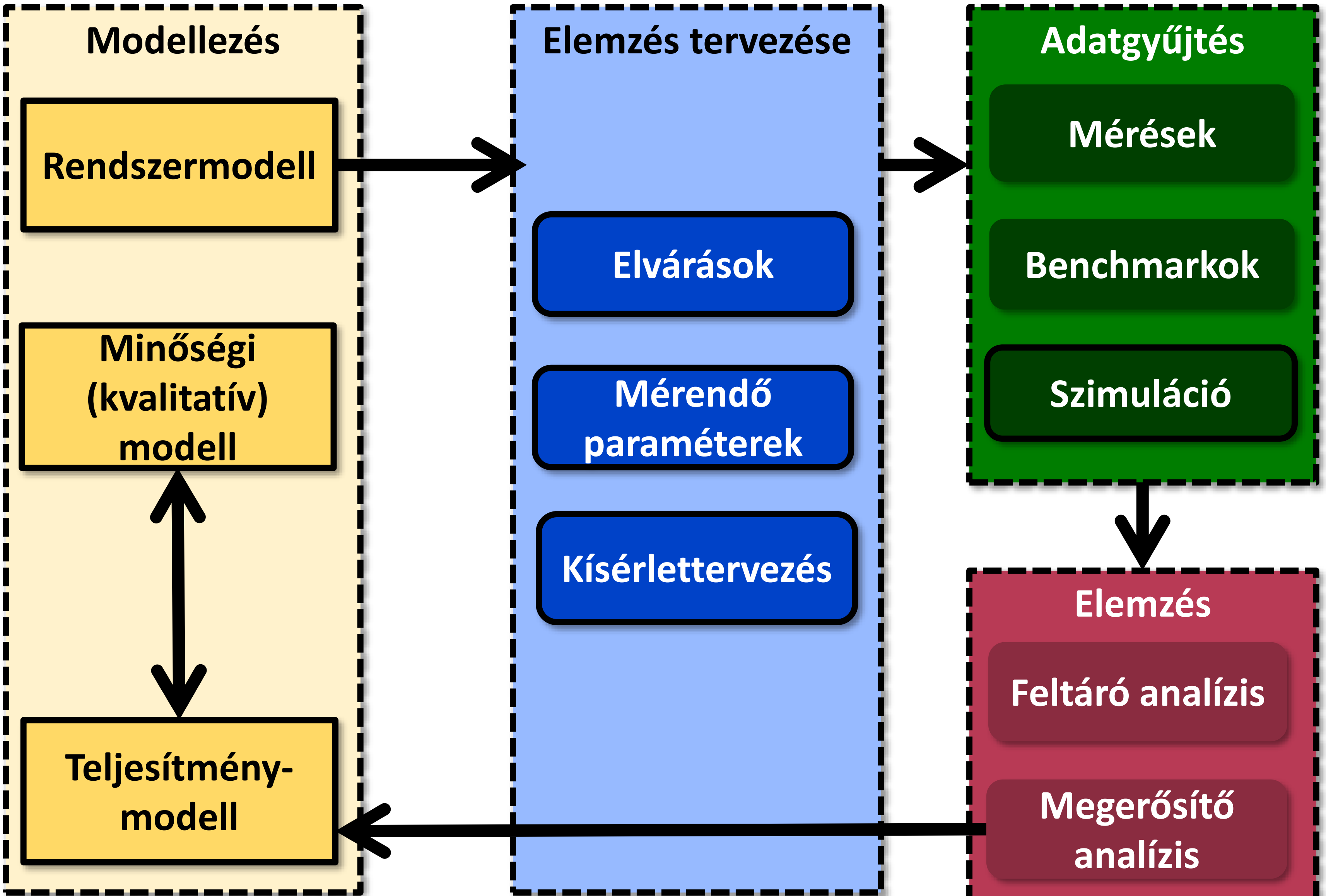
Terheléstesztelés

- „Szintetikus,” egyszerű terhelés
- Átbocsátóképesség felderítése
- Egy rendszer különböző verzióinak összehasonlítása
- Túlterhelés vizsgálata

Benchmarkolás

- Valós felhasználási eset alapján
- Összetettebb környezeti gerjesztések
- Különböző rendszerek objektív összehasonlítása
- Stabil működési tartományban

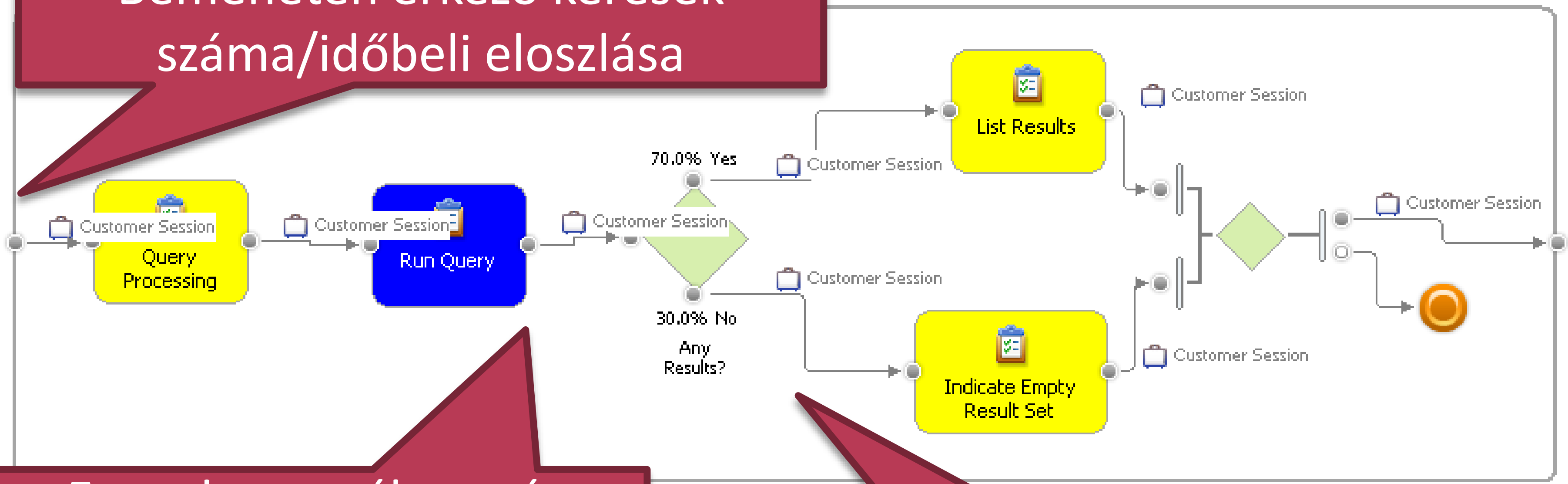
Rendszermodelltől a teljesítménymodellig



Alap kérdés

- Jól becsüljük meg a mennyiségi paramétereket?

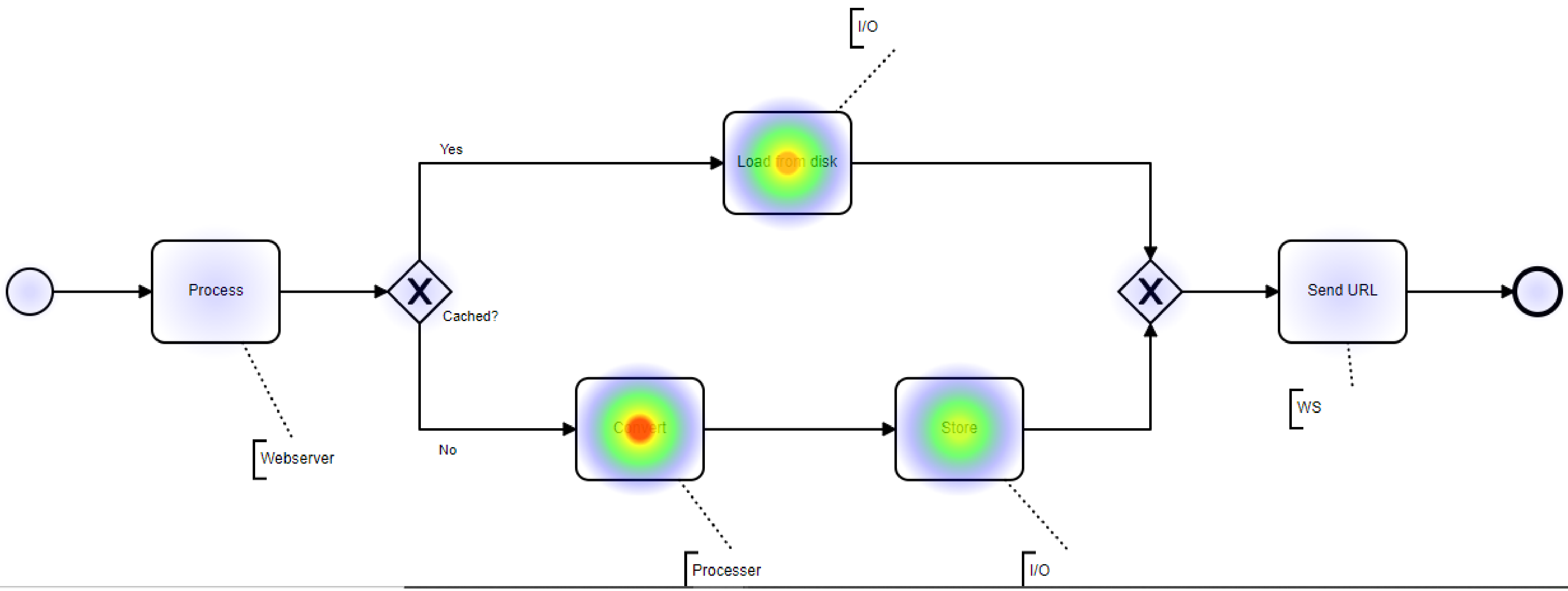
Bemeneten érkező kérések száma/időbeli eloszlása



Egy adott tevékenység végrehajtási ideje adott erőforráson

Várható értékkel közelített döntési valószínűség/gyakoriság

Példa eszköz (BIMP)



Adatok hihetősége

- Érzékenységvizsgálat
  - A modell **kimeneti** paraméterei mennyire érzékenyek a modell **bemeneti** paramétereinek változására
  - (erőforrások száma/kapacitása, felhasználók döntései) → (folyamat válaszideje, átbocsátása)
  - „parameter sweep”: egy paraméter vizsgálata adott tartományban

  → Melyik paramétert mennyire kell tudnunk becsülni?
- Ökölszabály: adatok hihetősége
  - mérés_bizonytalansága (szórás) ~ mérések_száma^2
  - (kellő mennyiségű adat esetén) → Val.szám.

MATEMATIKAI BECSLÉSEK: REGRESSZIÓS MÓDSZEREK

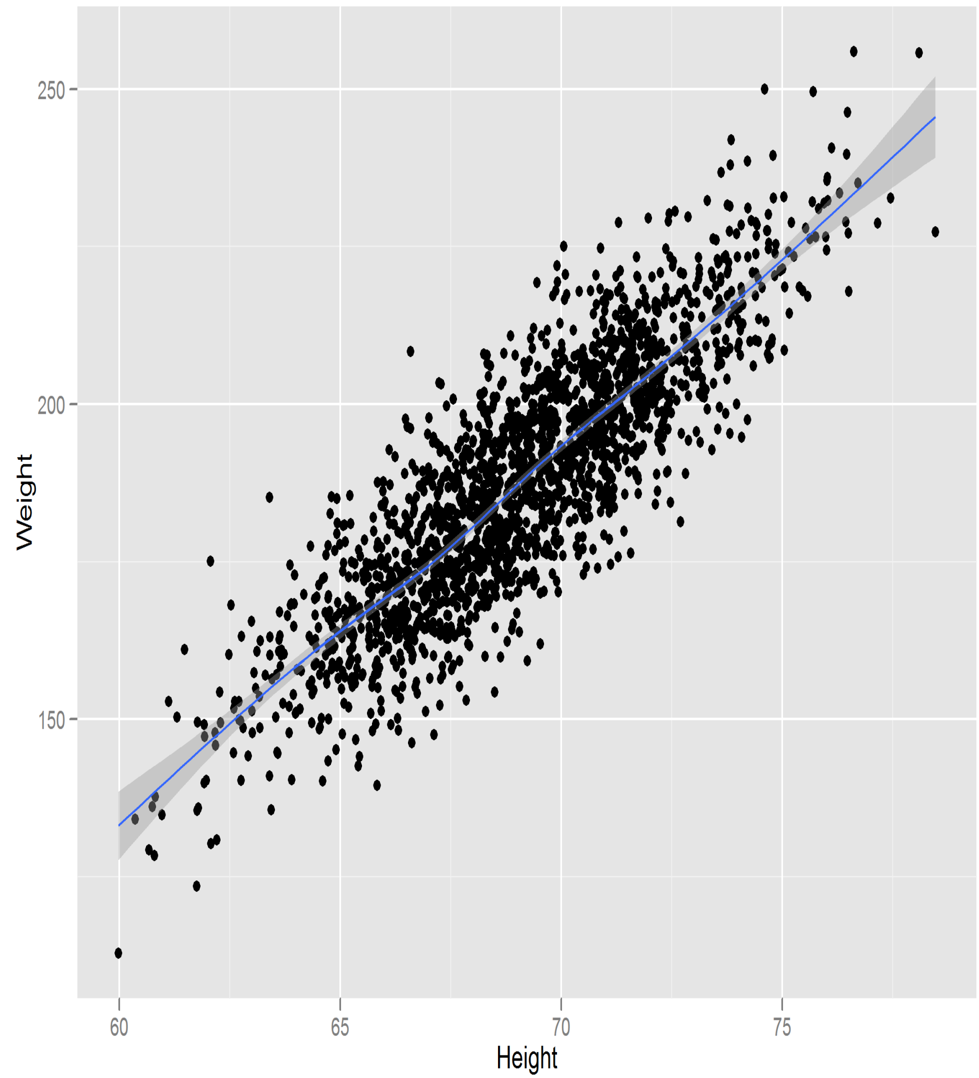
Feladat

- Adott sok változó, hosszú időre
- Szeretnénk bizonyos változók értékét közelíteni
  - Nehéz mérni
  - Nem áll rendelkezésre
- Szeretnénk becsülni/előrejelezni
  - Még nem történt meg, az idő függvényében közelítjük
  - Nem volt adott bemenet (pl. adott számú felhasználó)
  - Csak később lesz látható egy hatás (pl. válaszidő csak a kérések kiszolgálása közben nő meg a várakozás miatt)
- Mennyire bízhatunk meg a feltételezésekben/következtetésekben?

Regresszió

f függvény,

- bemenet:
  az attribútumok értéke,
- kimenet: megfigyelések legjobb közelítése
- „ökölszabály”
- Példa:
  testtömeg/magasság együttes eloszlás valójában egyenesre illeszthető



Regressziós módszerek

- Alapelv:

$$Y_t = f(\bullet) + \varepsilon_t$$

Véletlen változó

Közelítés

Hiba

Jósolt esemény

$$Y = f(X_1, X_2, \ldots, X_n)$$

Megfigyelhető változók

- Átlagos hiba (mean error)

$$ME = \frac{\sum_{t=1}^{n} (Y_t - F_t)}{n}$$

Becsült érték

Mért érték

Lineáris regresszió

- Egyszerű lin. függvény illesztése az adatokra
  - nem vár alapvető változást a rendszer viselkedésében

$$Y = a + bX$$

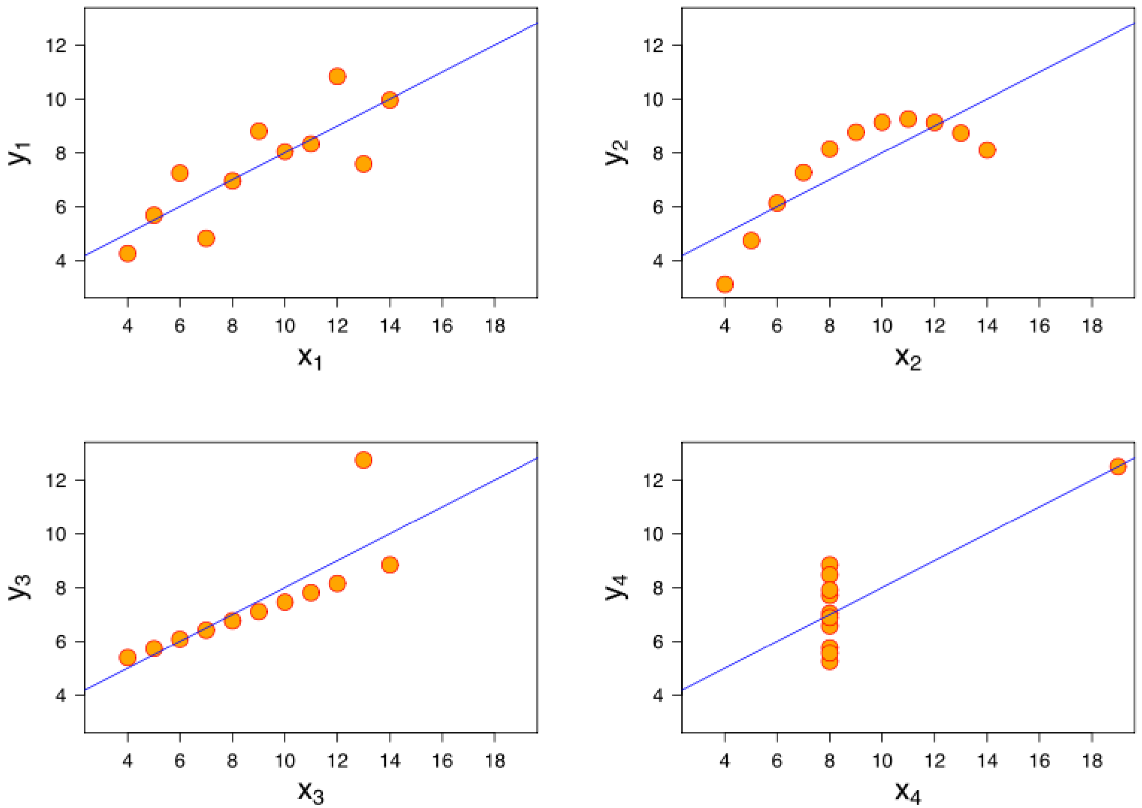
- Legkisebb négyzetek módszere
  - keressük azokat az *a,b* paramétereket (itt: a: eltolás, b: meredekség), amelyekre

$$SSE = \sum_{t=1}^{n} \varepsilon_t^{\,2} = \sum_{t=1}^{n} \left(Y_t - F_t\right)^2$$ minimális (Sum of Squared Errors)

- cél: $$\sum_{t=1}^{n} \left(Y_t - F_t\right)^2 = \sum_{t=1}^{n} \left[Y_t - \left(a + bX_t\right)\right]^2$$

Lineáris regresszió

- Legjobban illeszkedő egyenes
- **DE:** Anscombe's quartet
  - Minőségileg különböző adatok
  - Azonos regressziós egyenes



Lineáris regresszió (folyt.)

- Korrelációs együttható (négyzete)
  - a változó becsült és tényleges értékének kapcsolata
  - 0 és 1 közti érték
  - 0: nincs kapcsolat
  - 1: függvényszerű kapcsolat
  - R maga: -1 .. 1 (kapcsolat iránya)

$$R^2 = \frac{\sum_{t=1}^{n}\left(F_t - \overline{Y}\right)^2}{\sum_{t=1}^{n}\left(Y_t - \overline{Y}\right)^2}$$

- Példa: E-mail szolgáltatás, 8 hétig mérjük a csúcsterhelést.

Hét	1	2	3	4	5	6	7	8
Max. terhelés (email/perc)	420	410	437	467	448	460	507	514

Hogyan közelíthető a terhelés változása?

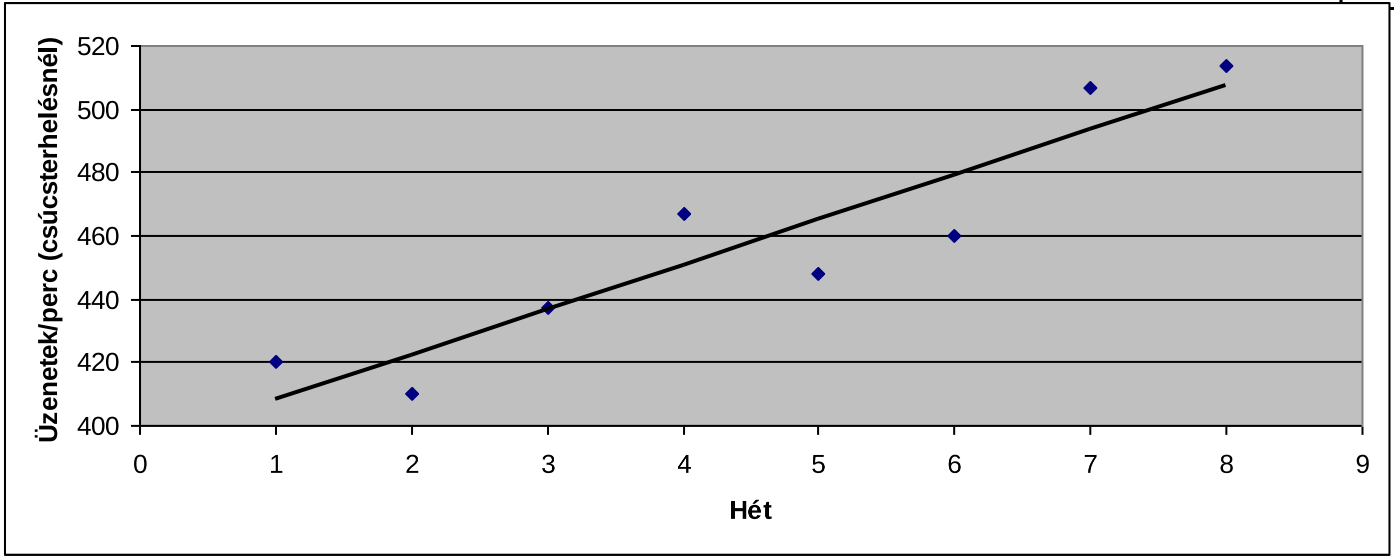
Mekkora a korrelációs együttható?

Lin. regresszió példa

A legkisebb négyzetek módszerével
Y=393.98+14.20X

Korrelációs együttható:
R^2=0.855

Mért	Jósolt terhelés
420	408.18
410	422.38
437	436.58
467	450.78
448	464.98
460	479.18
507	493.38
514	507.58
	521.78



Két változó kapcsolatának vizsgálata

- Tfh. lineáris kapcsolat van az egyszerre bejelentkezett felhasználók száma és az elküldött emailek közt. (pl logok alapján)

Bejelentkezett felh. átlagos száma (1 óra alatt)	2450	2765	2241	2860	3011	2907	3209
Átl. terhelés (kimenő+bejövő emailek/óra)	19257	20488	18152	21450	21077	20639	22142

- Lineáris regressziós közelítés a legkisebb négyzetek módszerével:

ÜzenetekSzáma = f(*BejelentkezettFelhasználók*)

Y=9480.48 + 3.95X, R^2=0.937 erős kapcsolat

Nemlineáris módszerek

- Exponceniális közelítés: $Y_t = a \times b^t$
  - jól illik a Web forgalom növekedéséhez
- Átalakítjuk a függvényt:

$$\log Y_t = \log a + t \log b$$

$$\log Yt = Y', \log a = a', \log b = b'$$

$$Y' = a' + b't$$

- Legkisebb négyzetek módszere használható
- Pl. adottak a legnagyobb mért terhelés értékei

Mennyi a várható legnagyobb terhelés az év végén?

Hónap	1	2	3	4	5	6	7	8	9	10
Max. kérések/sec (Y_t)	1035	1100	1160	1250	1350	1555	1770	1950	2210	2630
ln (Y_t)	6.942	7.003	7.056	7.13	7.207	7.349	7.478	7.575	7.7	7.874

Exp. terhelés példa

- Becslőfüggvény: $Y_t = a \times e^{bt}$
- Legkisebb négyzetek módszere a lineáris függvényre

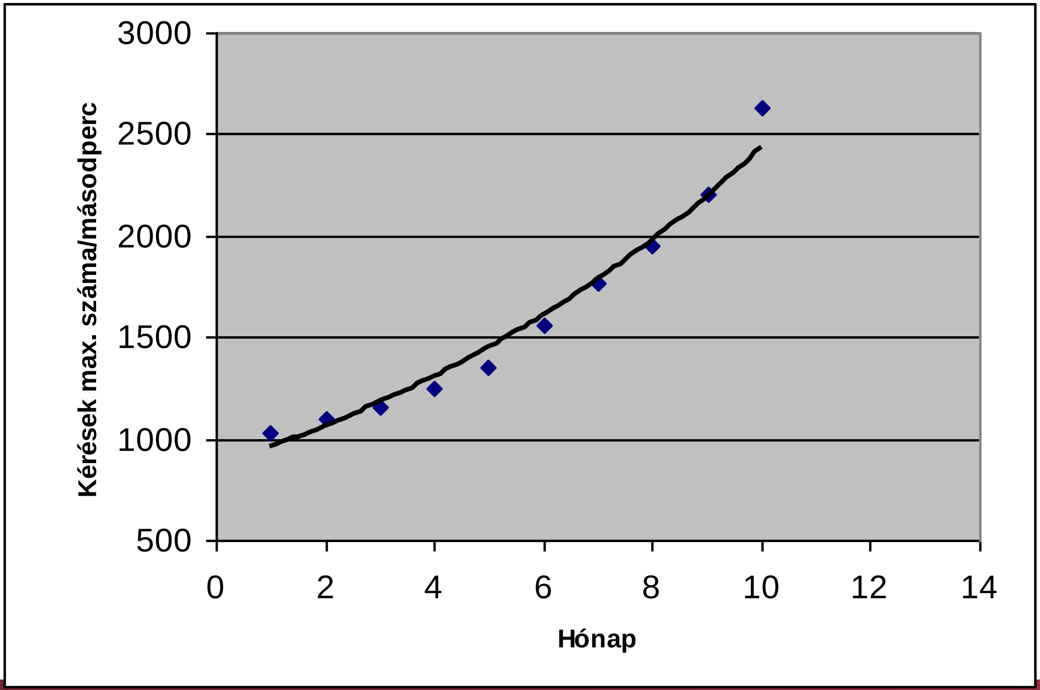
$$Y' = a' + b't, a' = 6.717, b' = 0.110, a = e^{a'}$$

- Eredmény:

$$Y_t = 826.33 \times e^{0.11t}$$

- 12. hónap:

Y_t=3093.3



Mozgó átlagok módszere

- Rövid távú előrejelzésre jó
- Egyszerre egy értéket ad meg
- A becsült érték az utolsó *n* érték átlaga

$$F_{t+1} = \frac{\sum_{i=t}^{t-n+1} Y_i}{n}$$

ahol $\boldsymbol{Y_t}$ a t. időpontban mért érték

$\boldsymbol{F_{t+1}}$ a becsült érték

n tipikusan 3 és 10 között van

(becslés hibája ne legyen túl nagy)

Exponenciális csúszóablak

- Egy értéket ad meg, az előző méréseket átlagolva
- Későbbi mérés (és mérési hiba) nagyobb súllyal
- Rövid távú előrejelzésre alkalmas
  - (miért exponenciális?)

$$F_{t+1} = F_t + \alpha\left(Y_t - F_t\right)$$

ahol $\boldsymbol{F_t}$:a t. időpontra becsült érték

$\boldsymbol{Y_t}$: t. időpontban mért érték

$\boldsymbol{Y_t - F_t}$: mérési hiba a t. periódusban

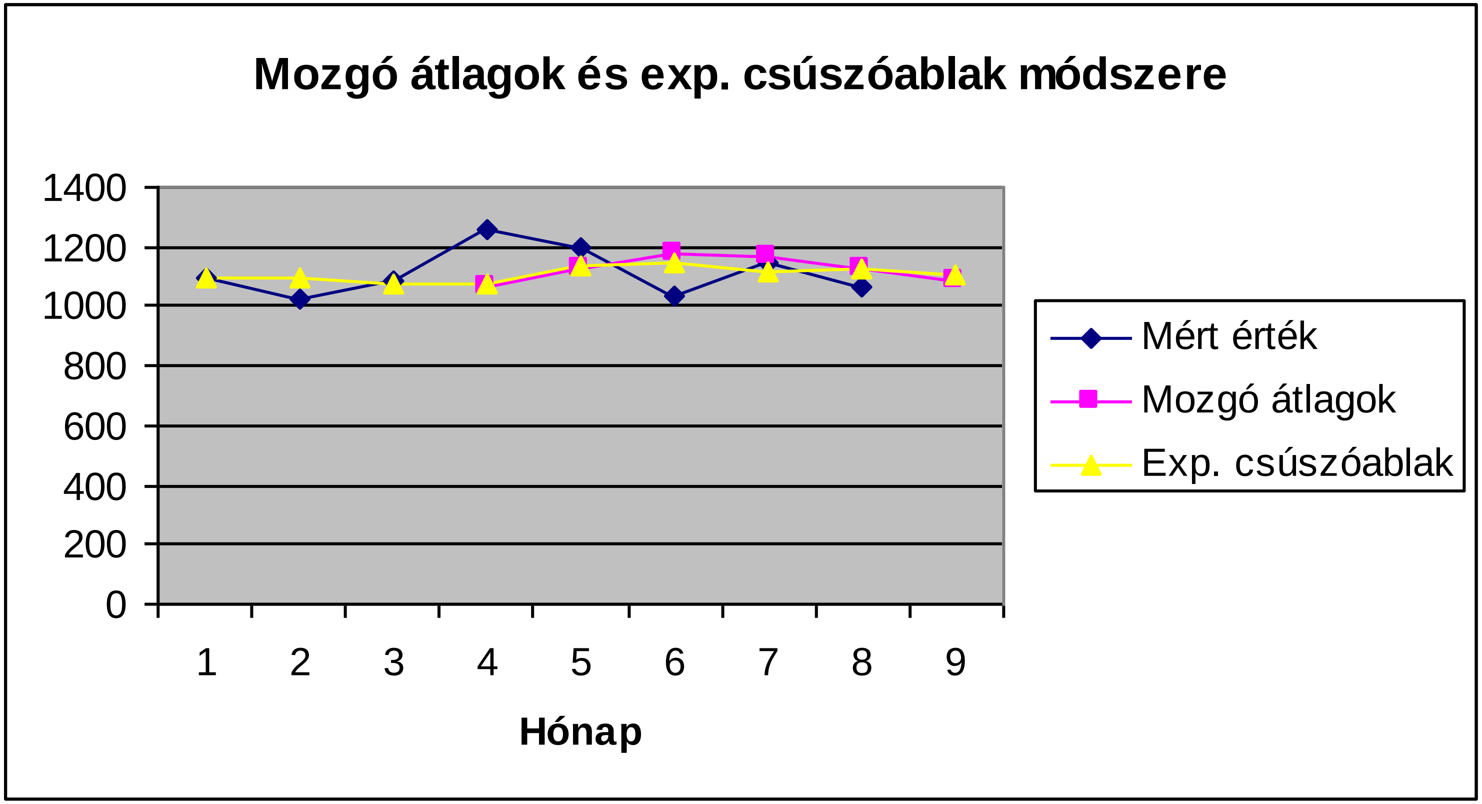
α : súlyozás ($0 \leq \alpha \leq 1$, gyakorlatban $0.05 \leq \alpha \leq 0.3$)

A két módszer összehasonlítása

- Adott sávszélesség igények, a két módszerel becsüljük a következő értéket

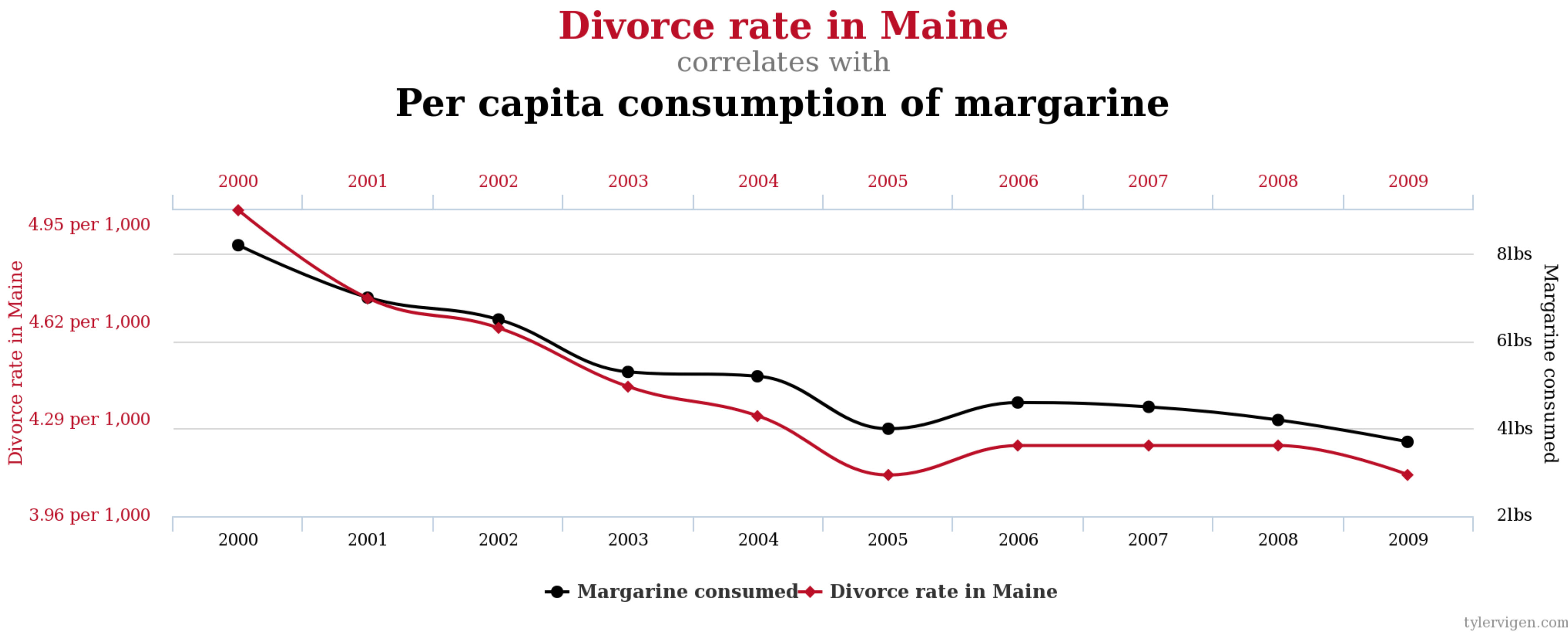
Hónap	Sávszélesség igény	Mozgó átlagok módszere (n=3)	Exp. csúszóablak (α = 0.3)
1	1100		1100.00
2	1020		1100.00
3	1090		1076.00
4	1255	1070.0000	1080.20
5	1195	1121.6667	1132.64
6	1039	1180.0000	1151.35
7	1145	1163.0000	1117.64
8	1066	1126.3333	1125.85
9		1083.3333	1107.90

A két módszer összehasonlítása



Mire figyeljünk?

Kauzalitás (oksági kapcsolat) != Korreláció (együttes előfordulás)



Informatikai példa: sok felhasználó → nagyobb kihasználtság ÉS magas válaszidő

A Little-törvény

A Zipf-törvény

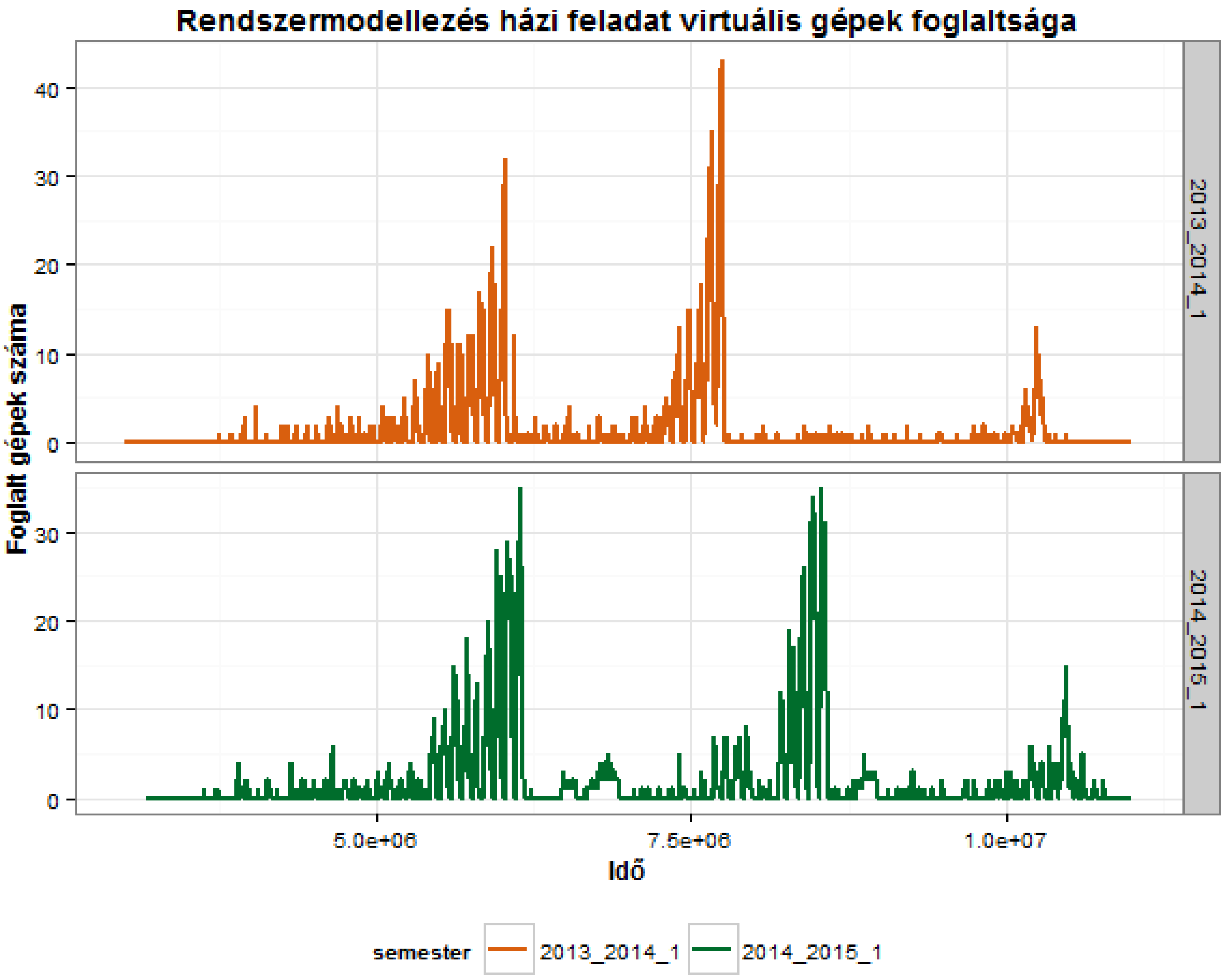
Terhelés változása

TERHELÉS VÁLTOZÁSA

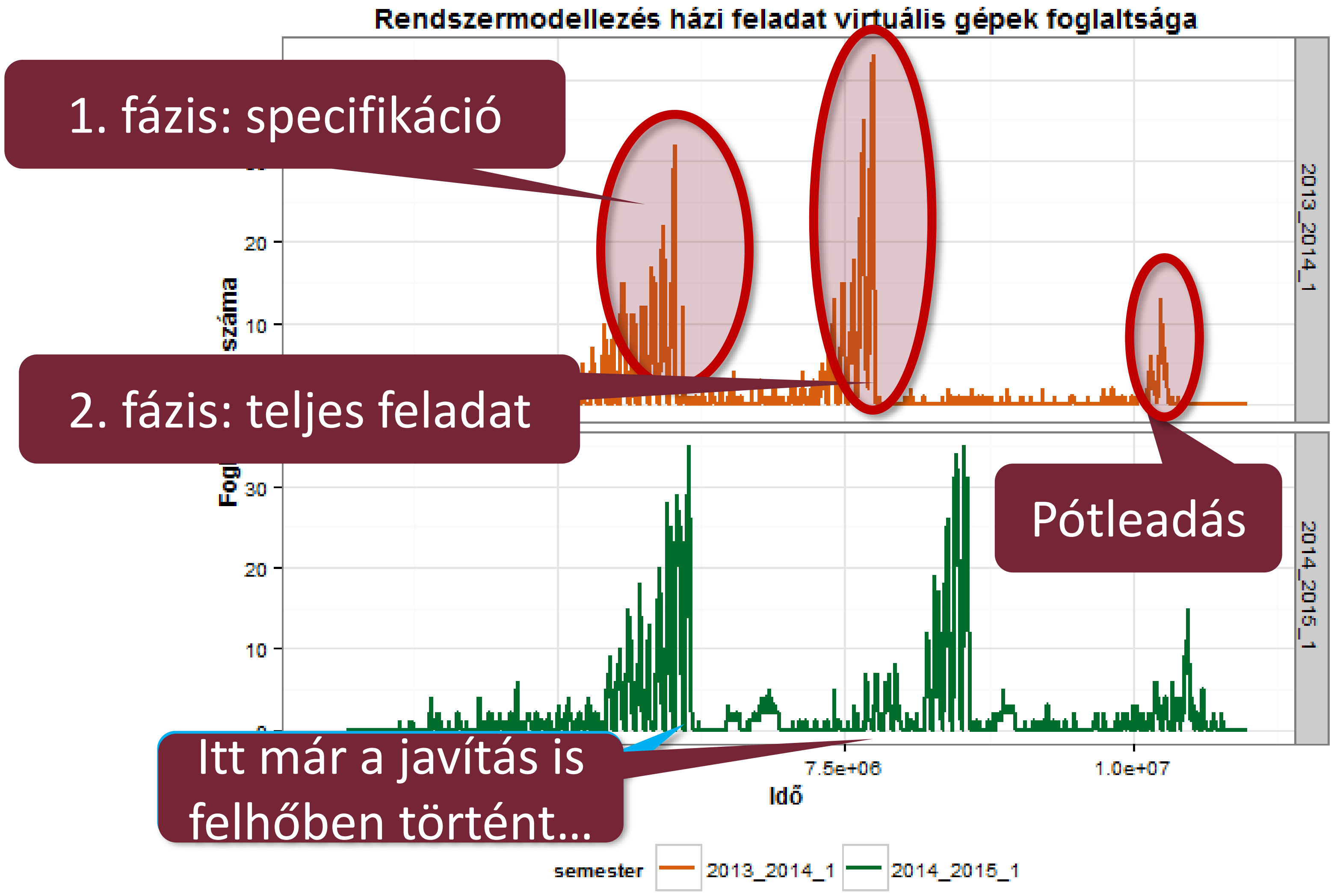
Milyen jellegű a terhelés?

- Eddig:
  - Átlagos értékekkel számoltunk
  - A rendszer viselkedését a *terhelés (intenzitás)* függvényében néztük
  - De: valójában nem (feltétlenül) előre kiszámíthatóan változik a terhelés
- Valójában
  - A rendszer viselkedése *időben* változik
  - Ennek műszaki hatásai vannak
    - Váltás feladatok közt, erőforrásfoglalás, stb. (pl. Operációs rendszerek)

Rendszermodellezés (7. félév) a felhőben



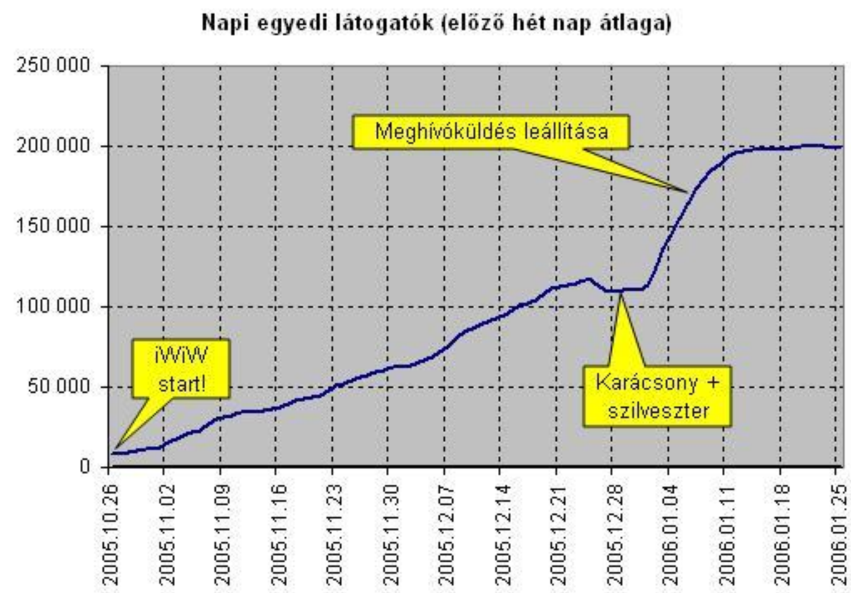
Rendszermodellezés (7. félév) a felhőben



Valós (történelmi) terhelés példa (iwiw)

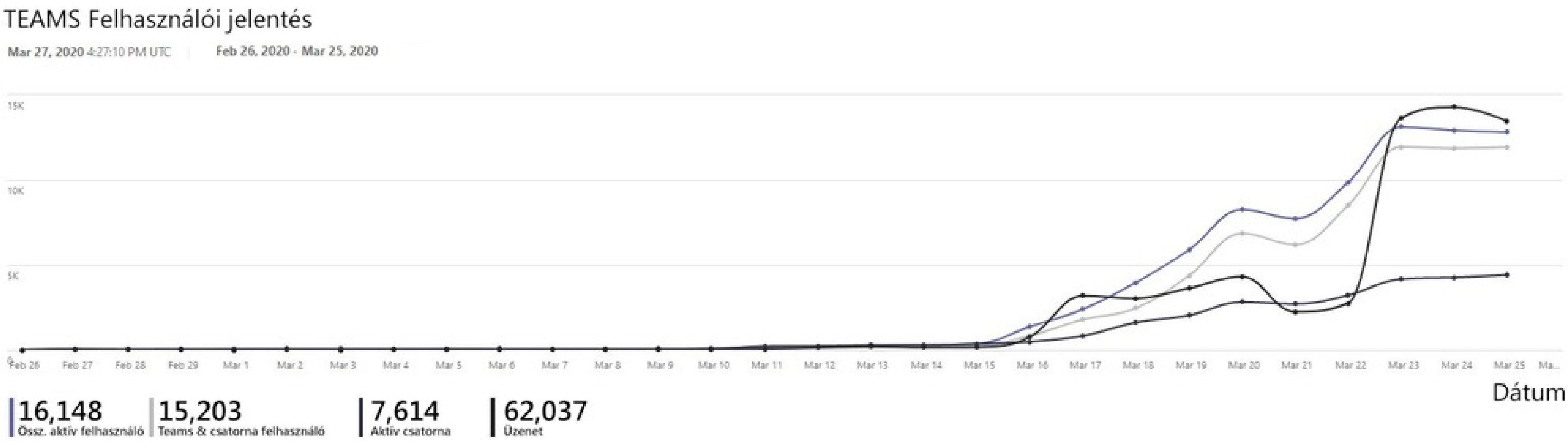
Szakaszokra becslőfüggvény illeszthető

- Lineáris, exponenciális, logaritmikus
- Regresszió, Bővebben pl. Valószínűségszámítás



Forrás: http://www.sg.hu/cikkek/42924/

Idei félév: Teams



http://www.vik.bme.hu/hir/2546-a-felhoben-folytatjuk

BENCHMARKING

Why benchmark?

Benchmarking - Definíció

- Wikipedia

*„In **computing**, a benchmark is the **act of running** a computer program, a set of programs, or other operations, in order to **assess the relative performance** of an object, normally by running a number of **standard tests** and trials against it.”*

A benchmarkolás

- egy **program** (programok, vagy más műveletek) **futtatása**
- **szabványos tesztekkel** vagy bemenetekkel
- egy objektum **relatív teljesítményének felmérése** érdekében

Benchmarking motiváció

- Fejlesztés alatt lévő rendszer
- Teljesítmény felmérése
  - Modell alapján becsült értékek
  - Megvalósított rendszer értékei „éles” helyzetben
- Tervezői és menedzsment döntések
  - Melyik részébe fektessünk több energiát (≈ pénzt)?
- Mi alapján döntsünk?
  - Jelenlegi erősségek és gyengeségek
- Mi számít erősségnek vagy gyengeségnek?
  - Mire képesek a hasonló rendszerek, versenytársak?

Elvárások

- Ismételhetőség (Repeatability)
  - Egy elemen többször megismételt mérés/művelet eredményeinek változékonysága (szóródási faktor)
- Reprodukálhatóság (Reproducibility)
  - A mérési rendszer változékonysága, amelyet a műveletek viselkedéseinek különbsége okoz
- Általánosított felhasználói eset
  - Átlag felhasználó számára értelmezhető legyen az eredmény

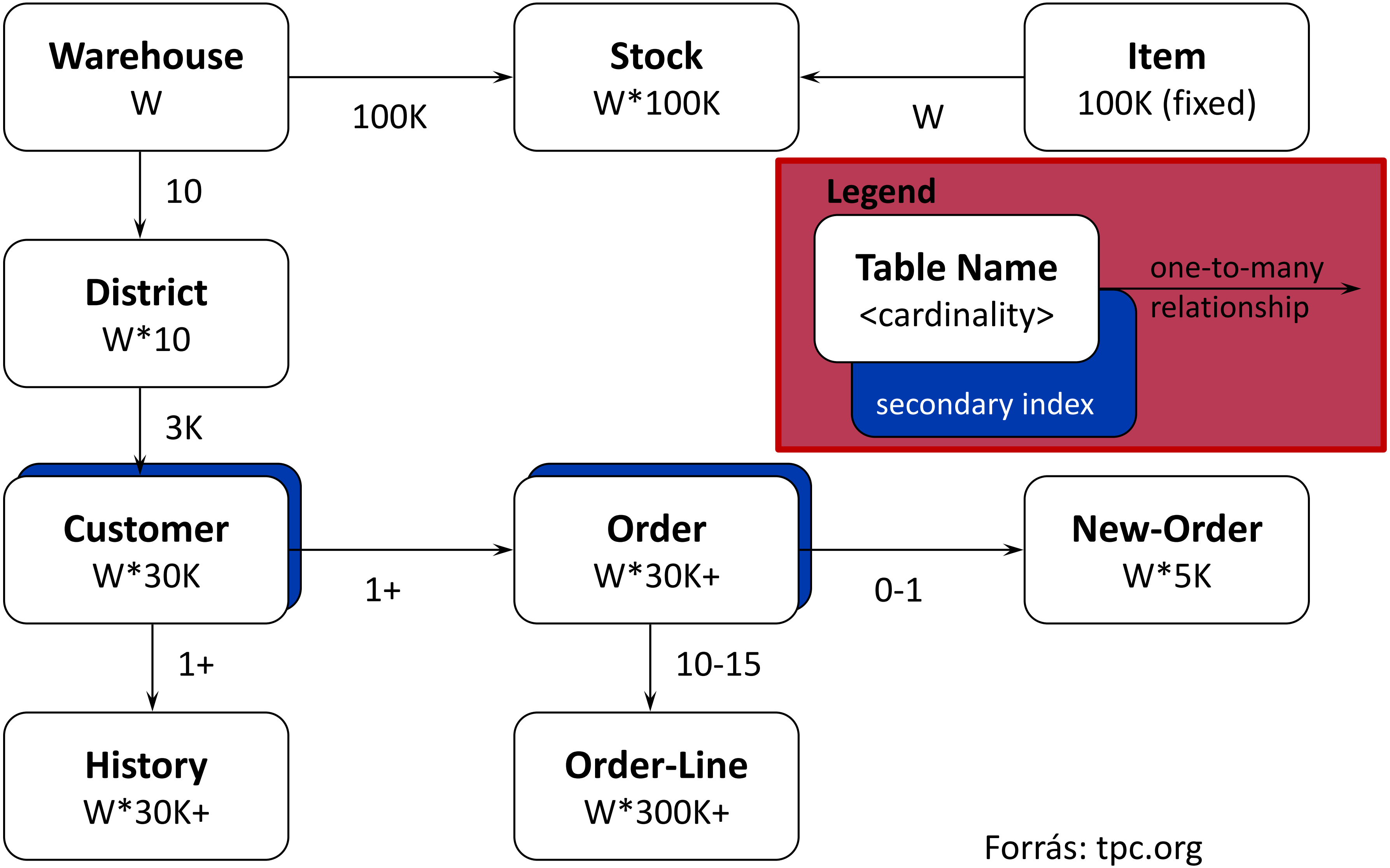
TPC-C

- OLTP benchmark
  - Online Transaction Processing
  - Szoftverek és hardverek OLTP teljesítményét méri
- Komplex benchmark
  - Több tranzakció típus
  - Összetett adatbázis séma
  - Széles skálán mozgó adathalmaz méret
- Mért metrikák
  - Áteresztőképesség: transactions per minute (*tpmC*)
  - Ár-teljesítmény arány (≈ hatékonyság): $/tpmC

TPC-C

- Mintaadatbázis
  - Nagykereskedelmi beszállító cég rendelései
  - 9 különböző tábla
- 5 féle tranzakció típus
  - Rendelés, fizetés, szállítás, rendelés státusza, raktár állapota
  - Frissítés, beszúrás, törlés, megszakítás
- ACID tranzakciók
  - Atomicity, Consistency, Isolation, Durability
- Változatos felhasználói kérések szimulációja
  - Tranzakció típusok kezdeményezése adott valószínűséggel
- Időkorlát az egyes tranzakció típusokra

TPC-C adatbázis séma



Forrás: tpc.org

KIÉRTÉKELÉS

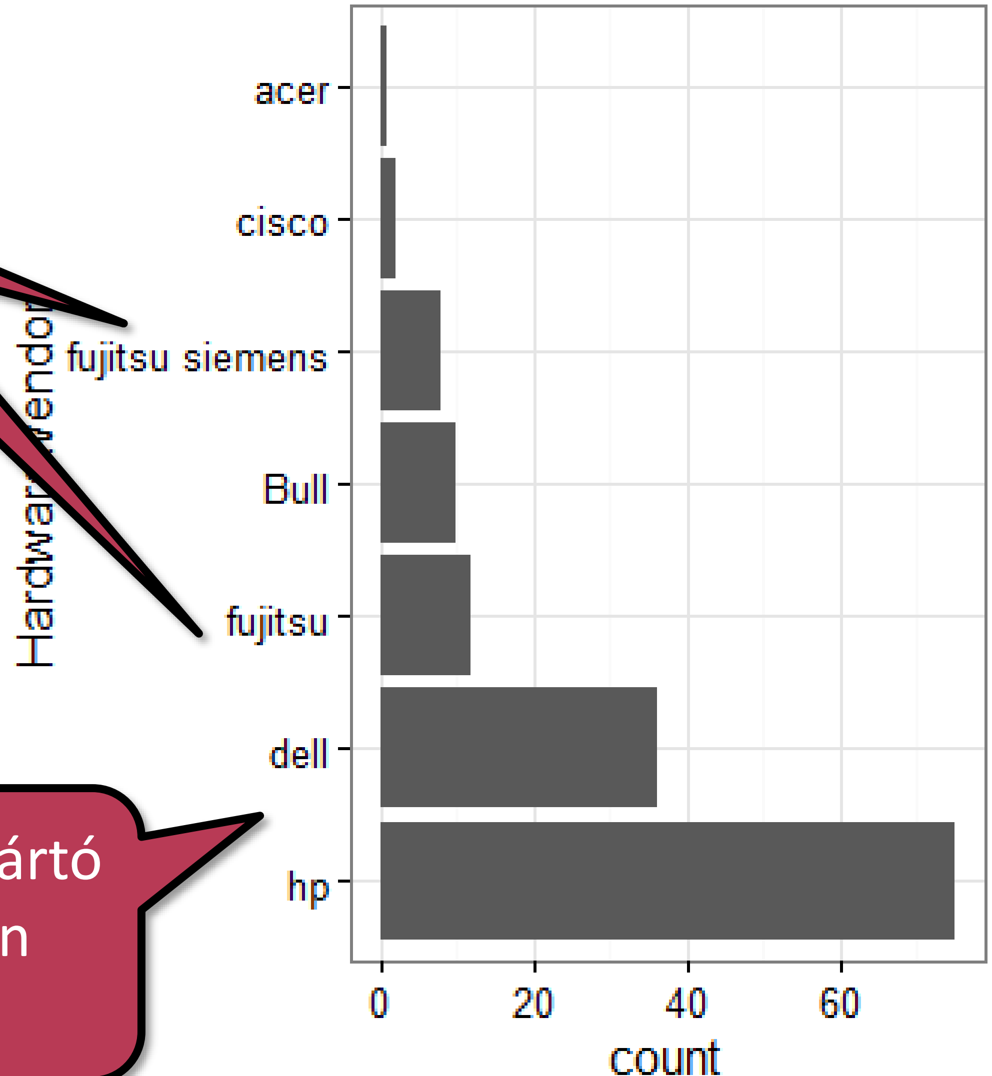
Adattisztítás

- Pénznem egységesítés (EUR, USD, AUD, JPY)
- Tizedesjegy, tízes elválasztó ellenőrzés

Hardware Vendor	System	tpmC	Price/tpmC	Watts/KtpmC	System Availability	Database	Operating System	TP Monitor	Date Submitted
acer	Altos R710 12GB/3.6GHz/2P C/S with 4 Altos R710	66,543	12.42 AUD	NR	12/01/05	Microsoft SQL Server 2000 Enterprise Edition SP3	Microsoft Windows Server 2003 Enterprise Edition	Microsoft COM+	01/12/06
Bull	Bull Escala PL6460R	6,085,166	2.81 USD	NR	12/15/08	IBM DB2 9.5	IBM AIX 5L V5.3	Microsoft COM+	06/15/08
Bull	Bull Escala PL860R	629,159	2.49 USD	NR	06/11/08	IBM DB2 9.5 Enterprise Edition	IBM AIX 5L V5.3	Microsoft COM+	06/11/08

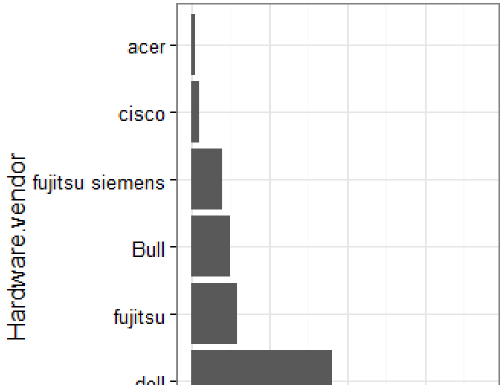
Milyen rendszereket vizsgálunk?

Egységesítsünk?

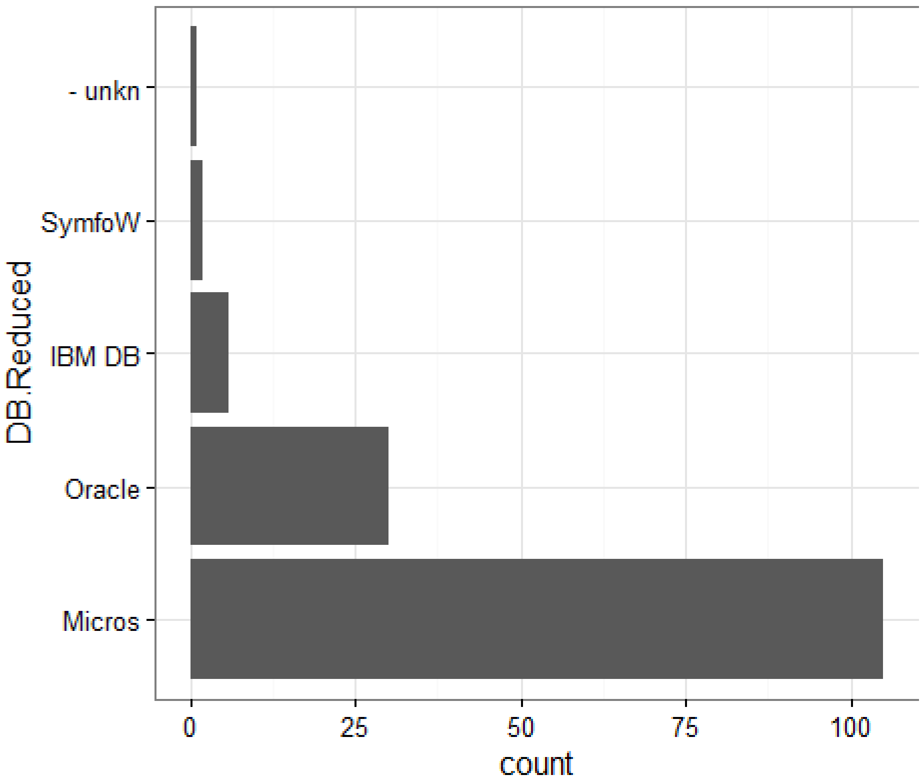


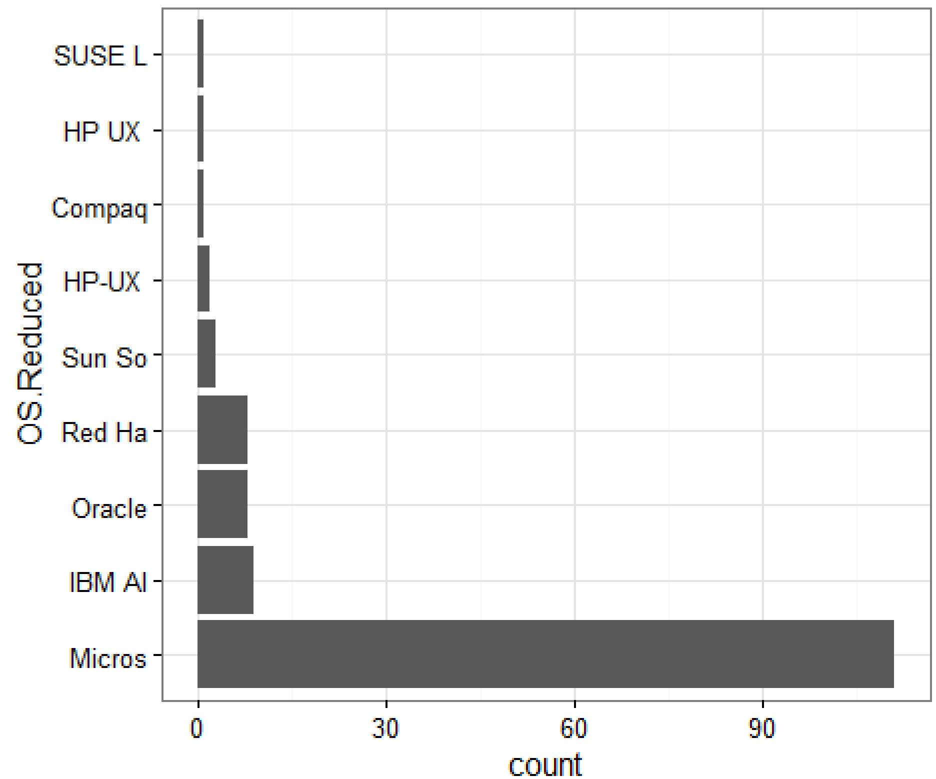
A két leggyakoribb gyártó az esetek 77%-ában szerepel

Milyen rendszereket vizsgálunk?

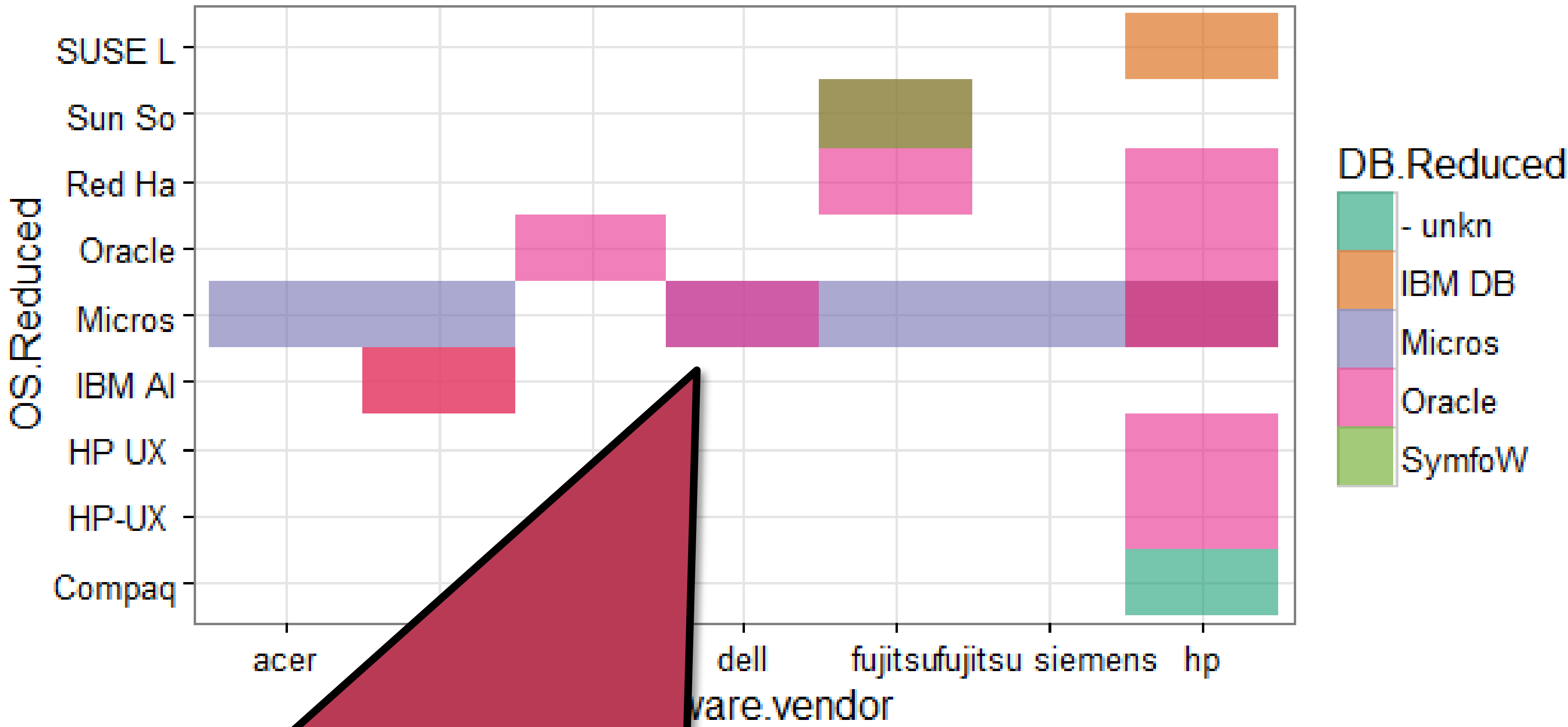


Típus szerinti összevonás?



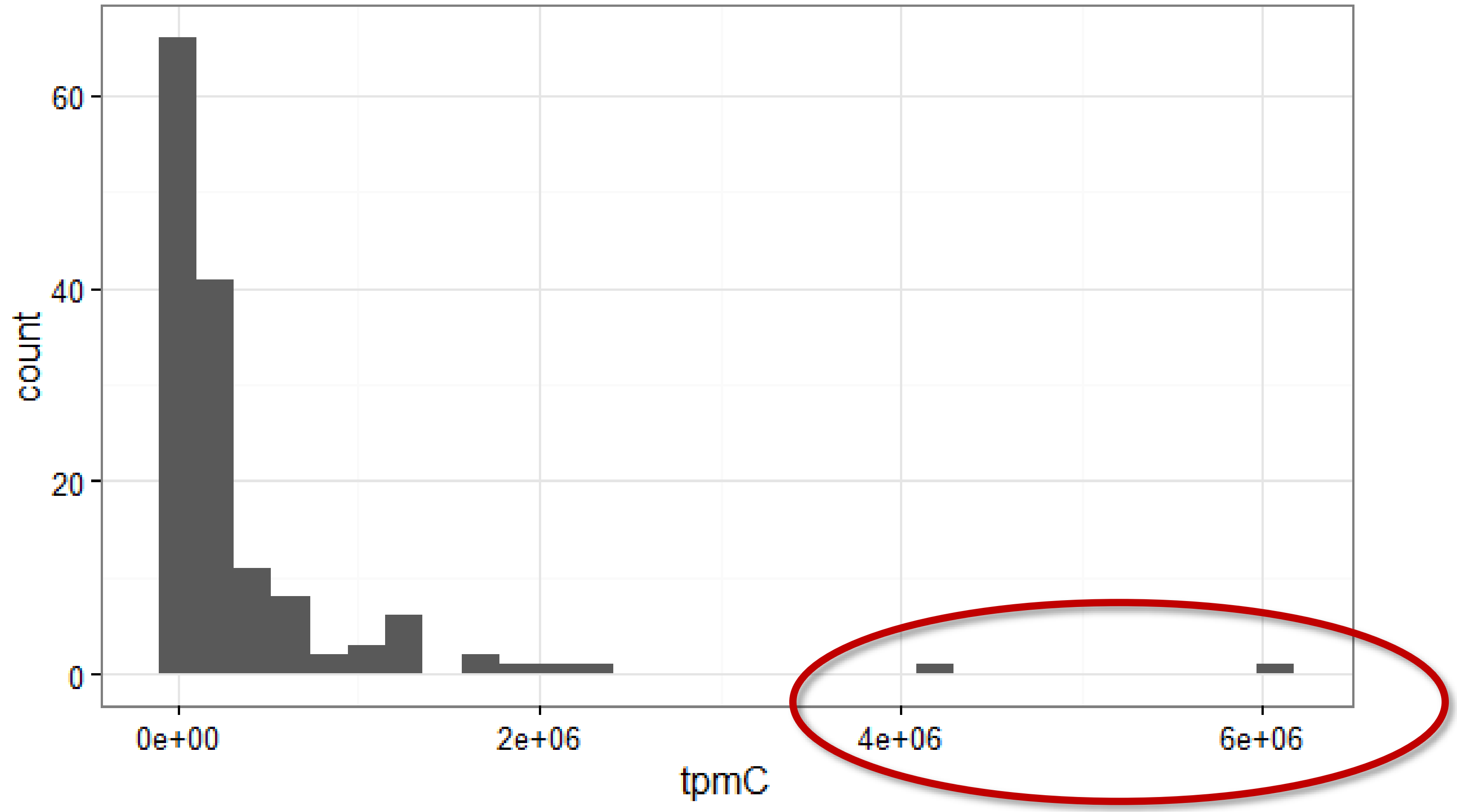


Milyen rendszereket vizsgálunk?



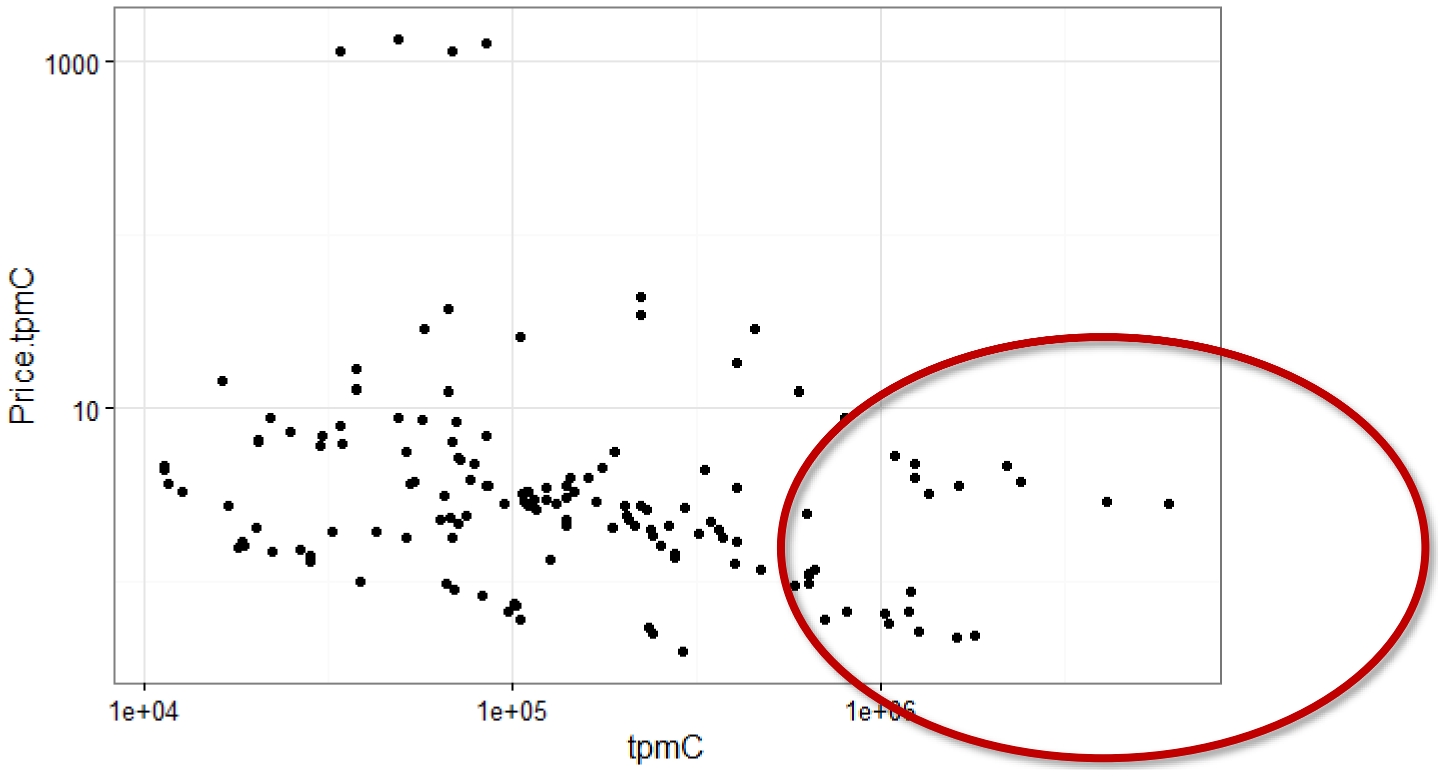
Általában többféle adatbázis többféle operációs rendszeren fut

Benchmark eredmények



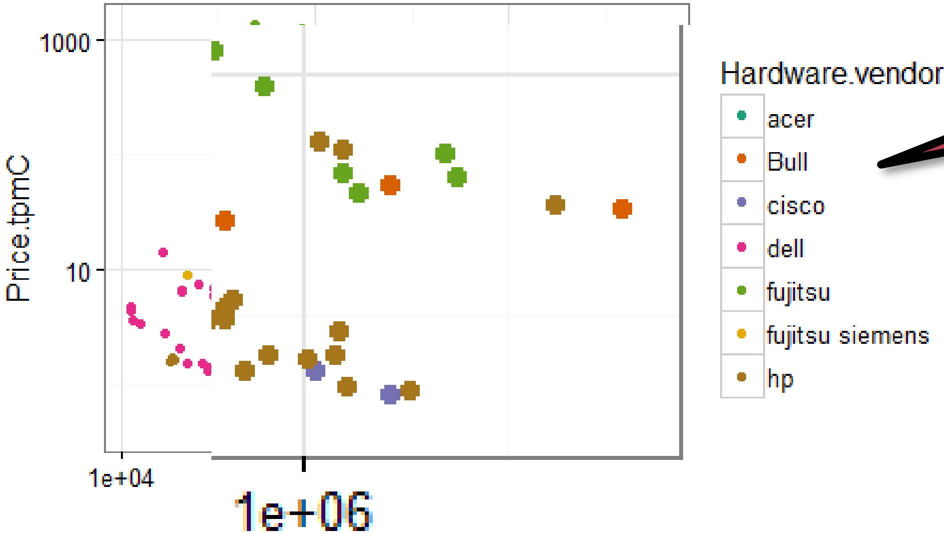
Milyen az ár/érték arányuk?

Benchmark eredmények

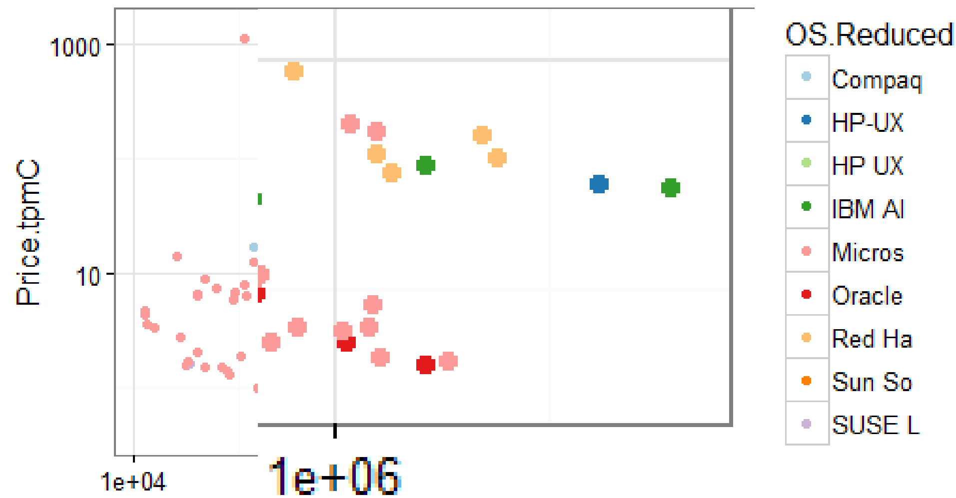


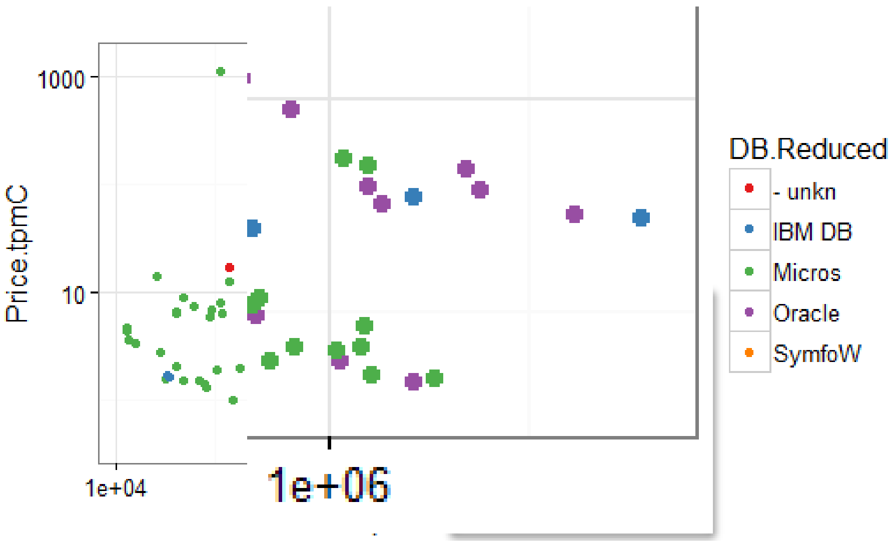
Logaritmikus skála?

Benchmark eredmények



Az élmezőny elég változatos

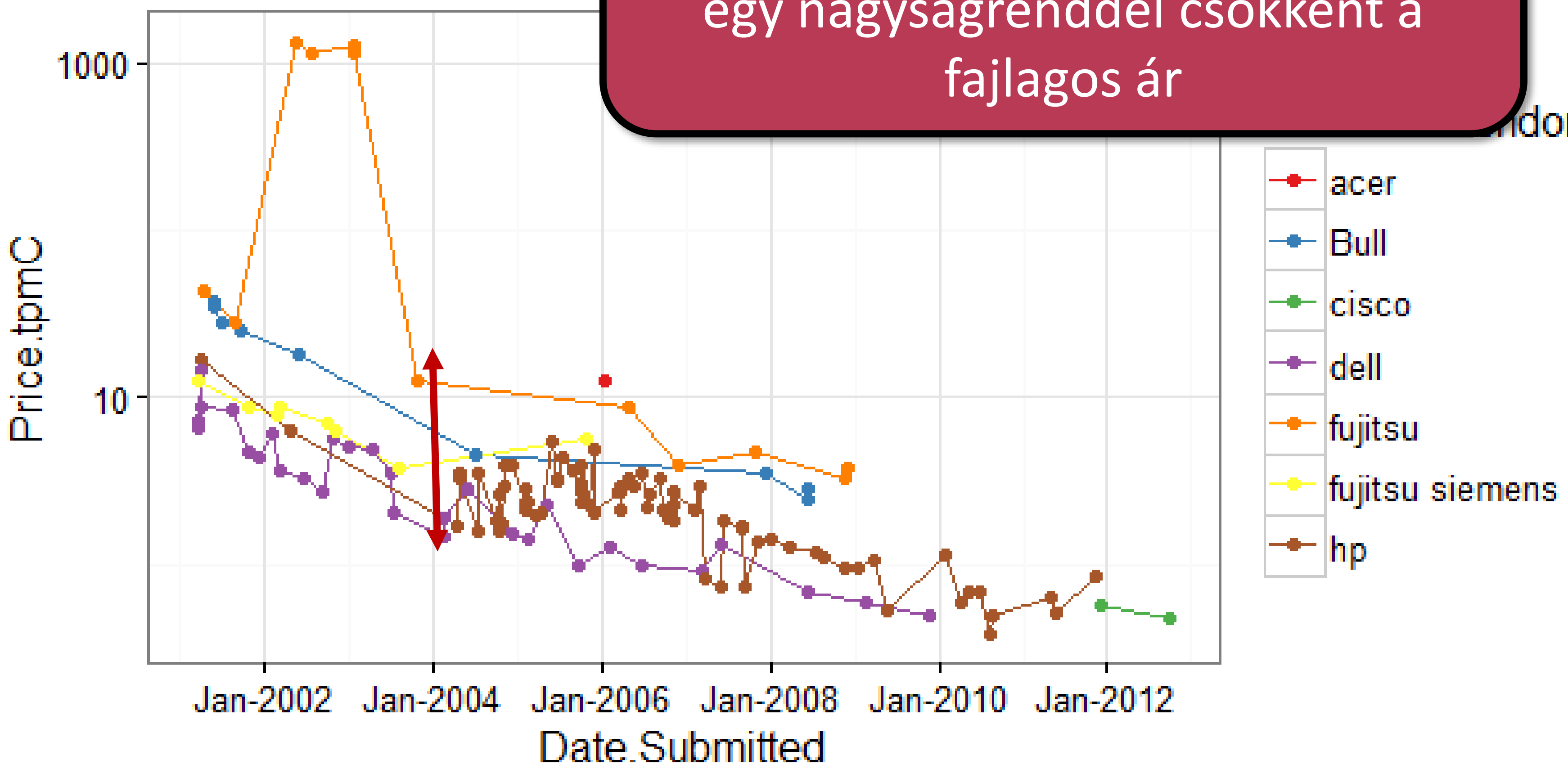




Nincs legjobb OS és DB konfiguráció sem

Időbeli változások?

A legtöbb gyártónál
10 év alatt
egy nagyságrenddel csökkent a
fajlagos ár



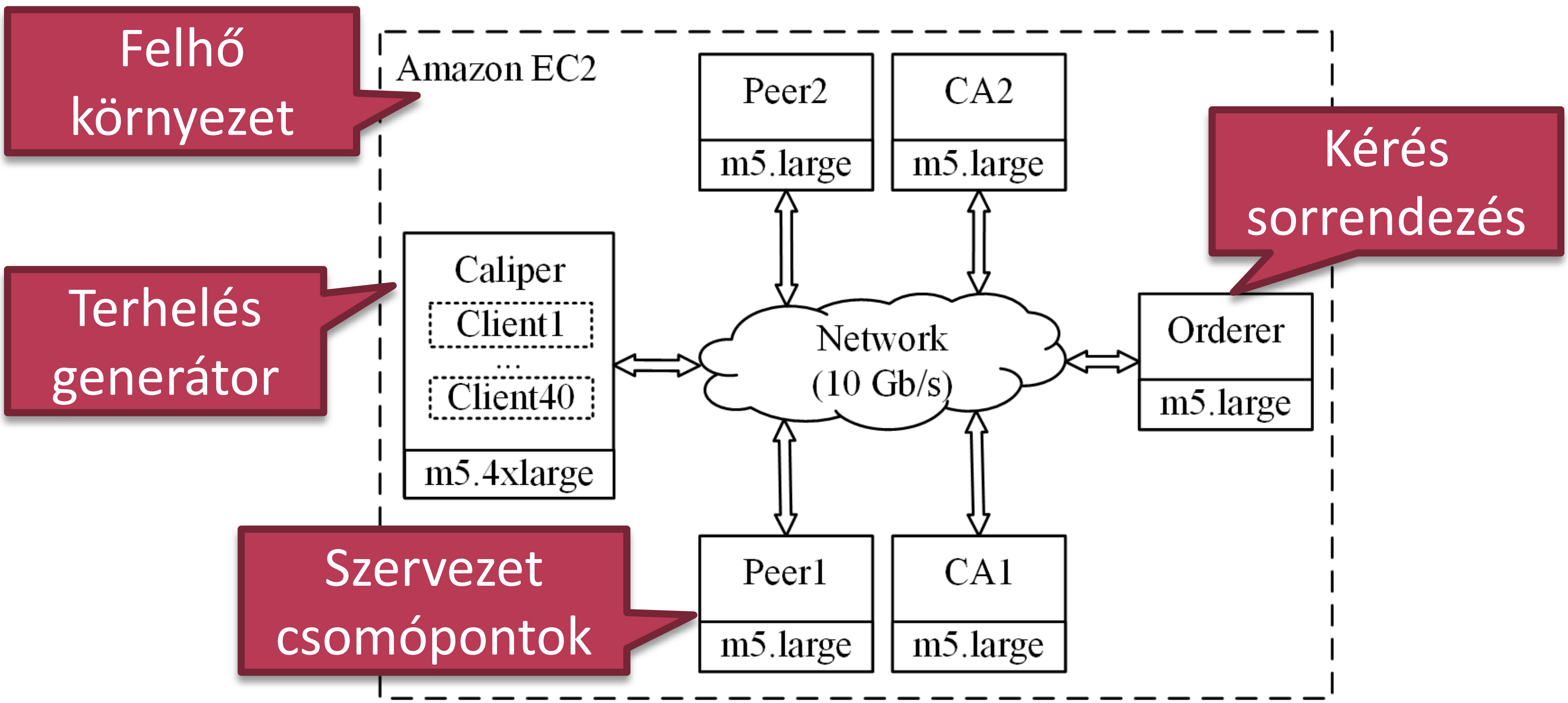
TELJESÍTMÉNYTESZTELÉS

Esettanulmány: Hyperledger Fabric

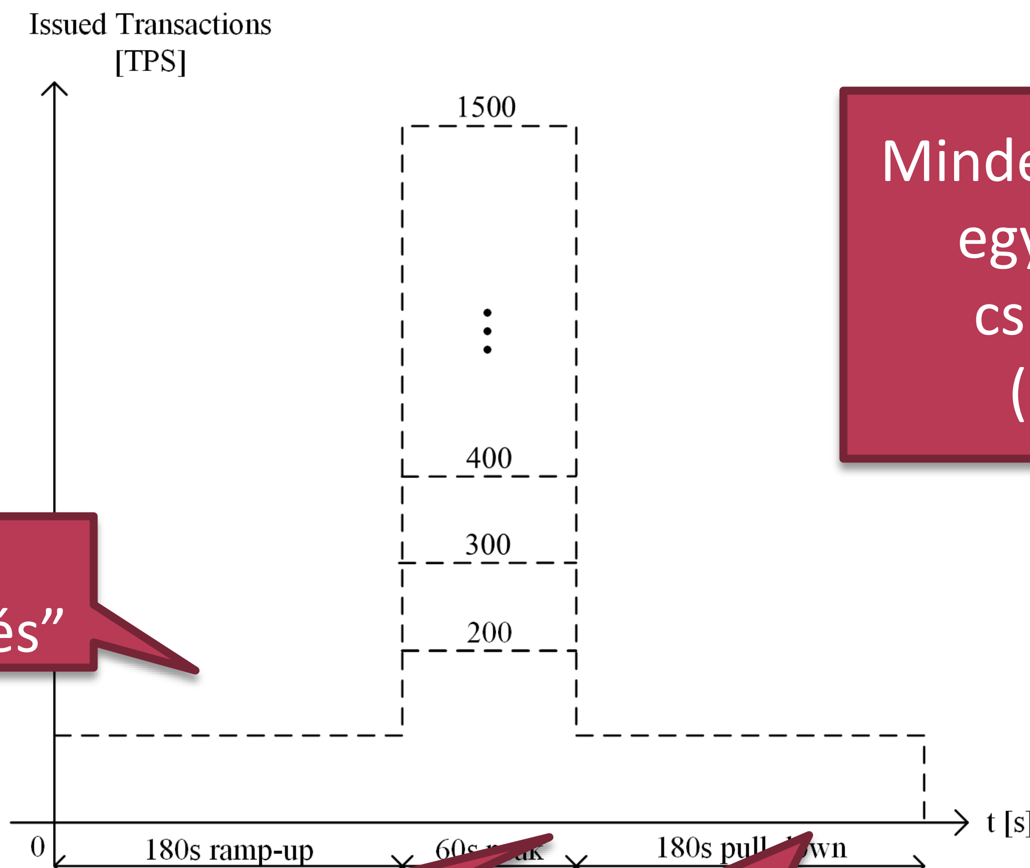
https://www.hyperledger.org/projects/fabric

Hyperledger Fabric

- Konzorciális elosztott főkönyv technológia (DLT)
- Komplex infrastruktúra
  - Szervezetek, főkönyvek, telepített okos szerződések



Terhelésprofil



Minden mérés során egyre nagyobb csúcsterhelés (14 mérés)

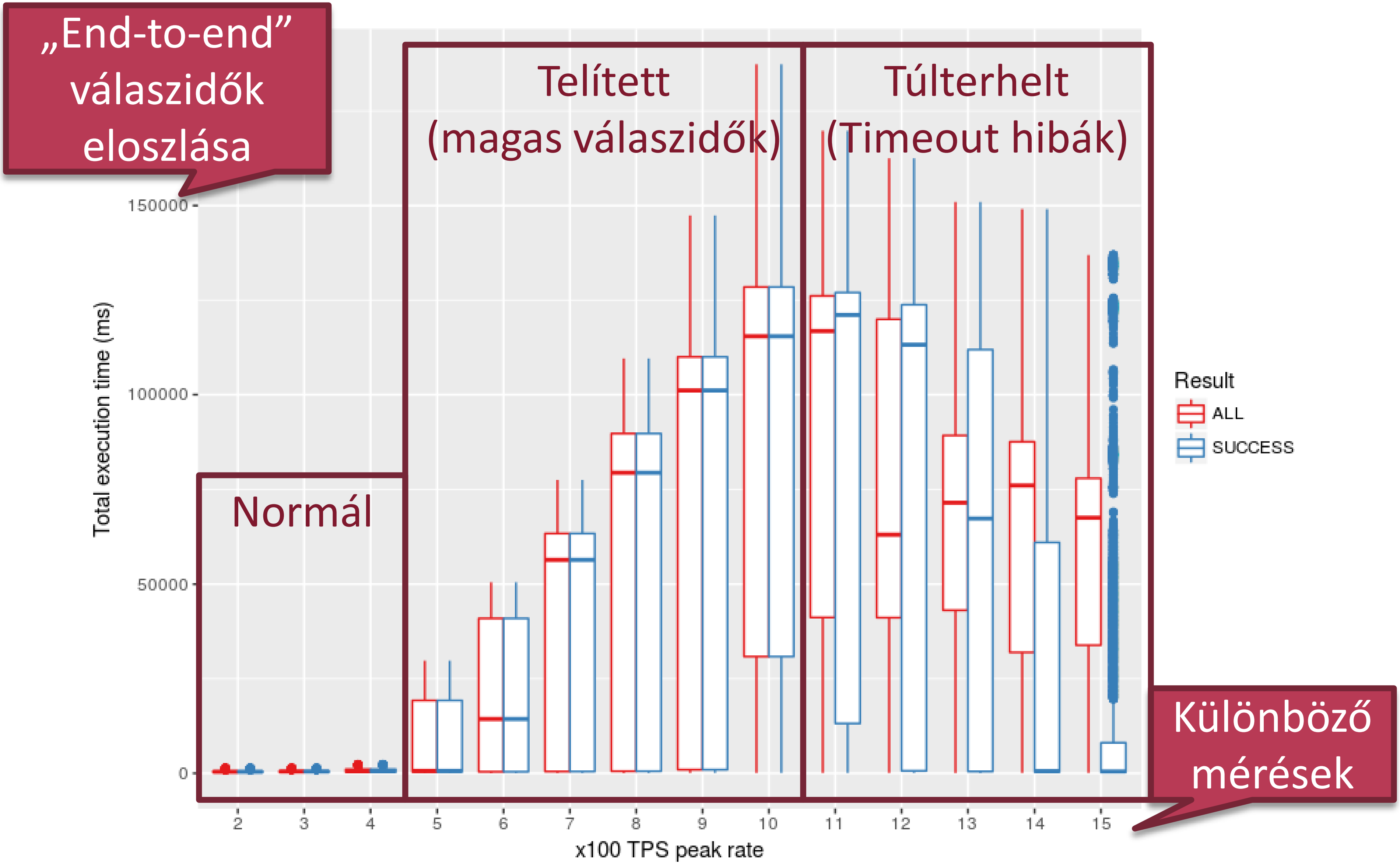
3 perc „bemelegítés”

1 perc intenzív terhelés

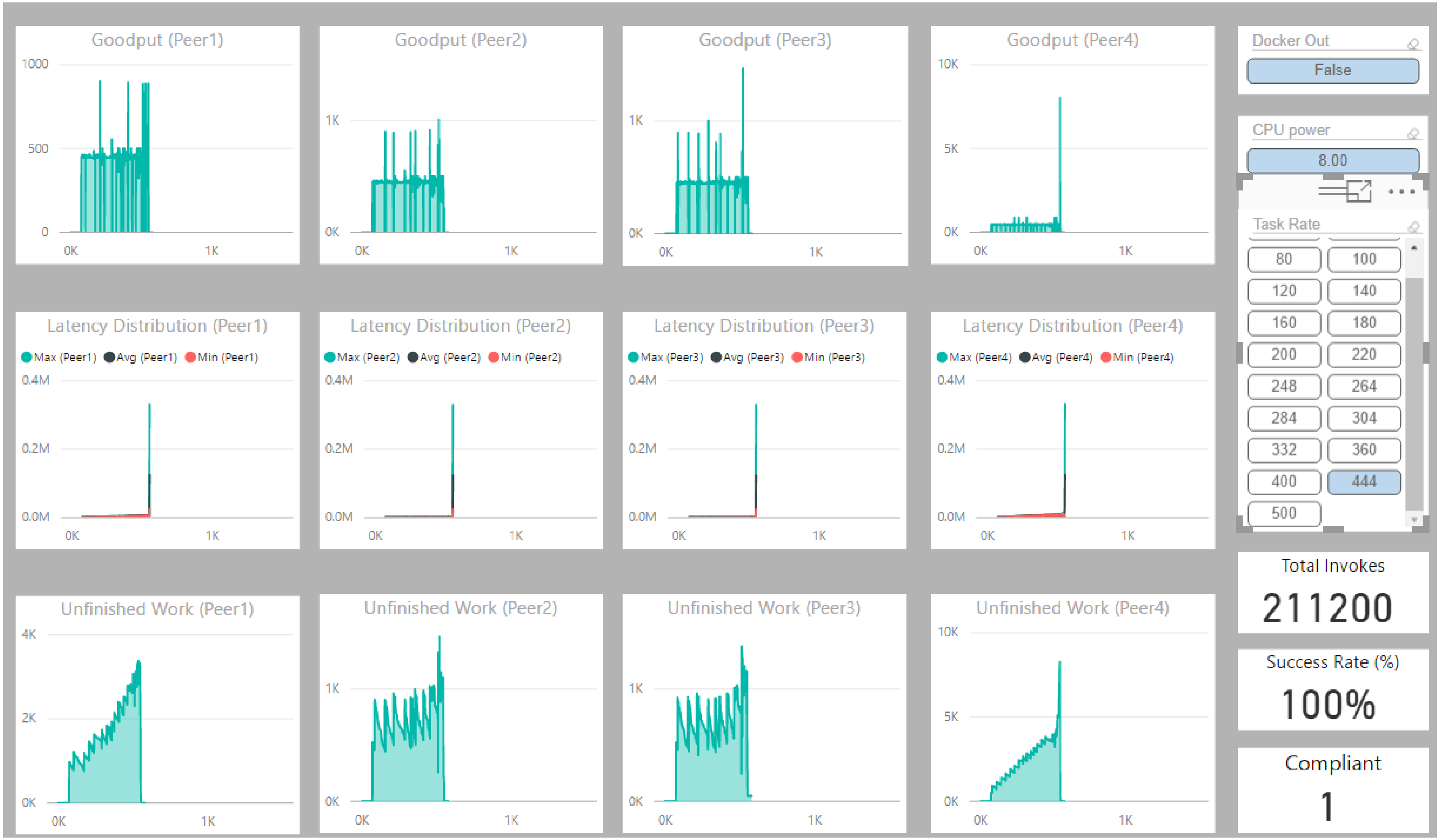
3 perc „kifutási” idő

7 perces terhelés

Válaszidők eloszlása



A motorháztető alatt..



KÍSÉRLETTERVEZÉS ALAPFOGALMAK

Kísérlettervezés

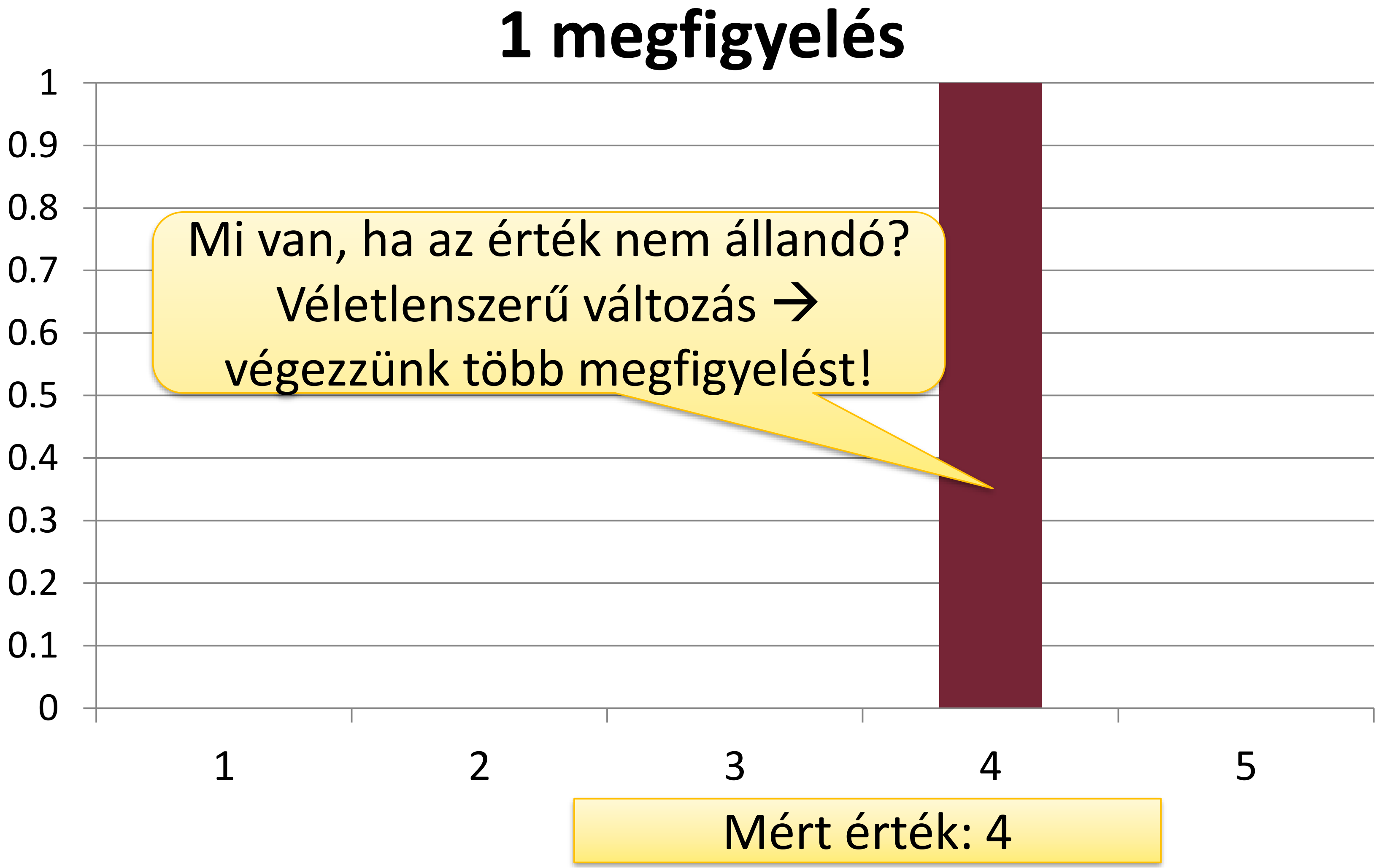
- Cél: a modell paraméterezése a valóság alapján
  - Vagy absztrakt modell a konkrét modell alapján
- Információt **kísérlet**ek révén szerzünk
  - Pontosan mit szeretnénk tudni?
  - Ehhez milyen megfigyelést, hányszor kell elvégezni?
  - A kapott eredményekből mire lehet következtetni?
- (statisztikai) **kísérlettervezés** (Design of Experiment, DOE)
  - hatékony eljárás a kísérletek tervezésére és elemzésére
  - valós és objektív konklúziók levonásához
- Kísérletterv: még a kísérlet elvégzése előtt

Kísérlettervezés

- Mire jó?
  - Alternatívák közötti választás
  - Érzékeny paraméterek, kulcsfaktorok
  - Megfelelő célérték, változékonyság csökkentése
  - Robosztussá tétel
- Fontos:
  - Világos cél, egyértelmű eredmények
  - Kis méret, alacsony költség
  - Valós viszonyok

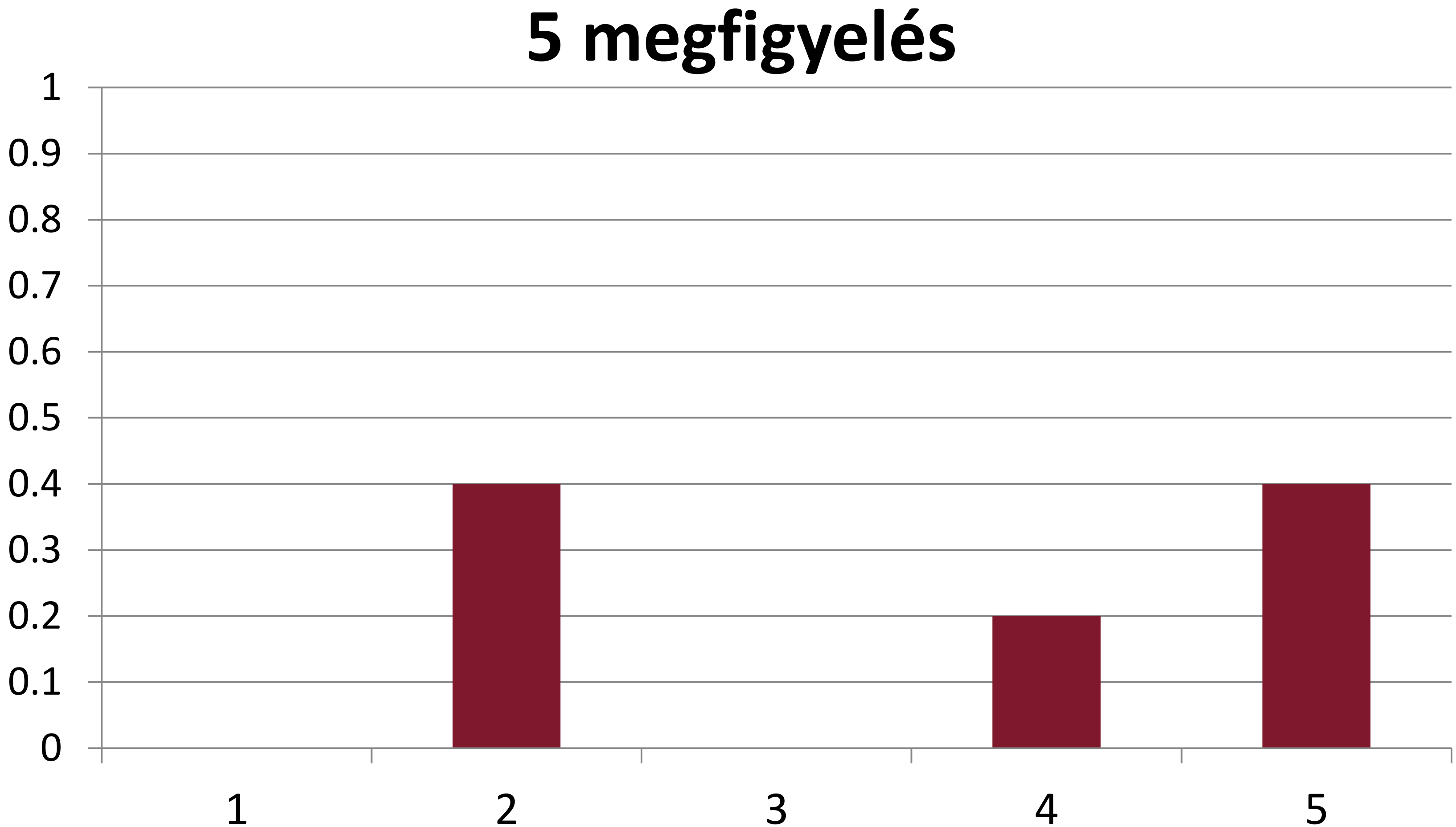
Számszerű jellemzők mérése

- Mérjük meg egy jellemző értékét! (1..5)



Számszerű jellemzők mérése

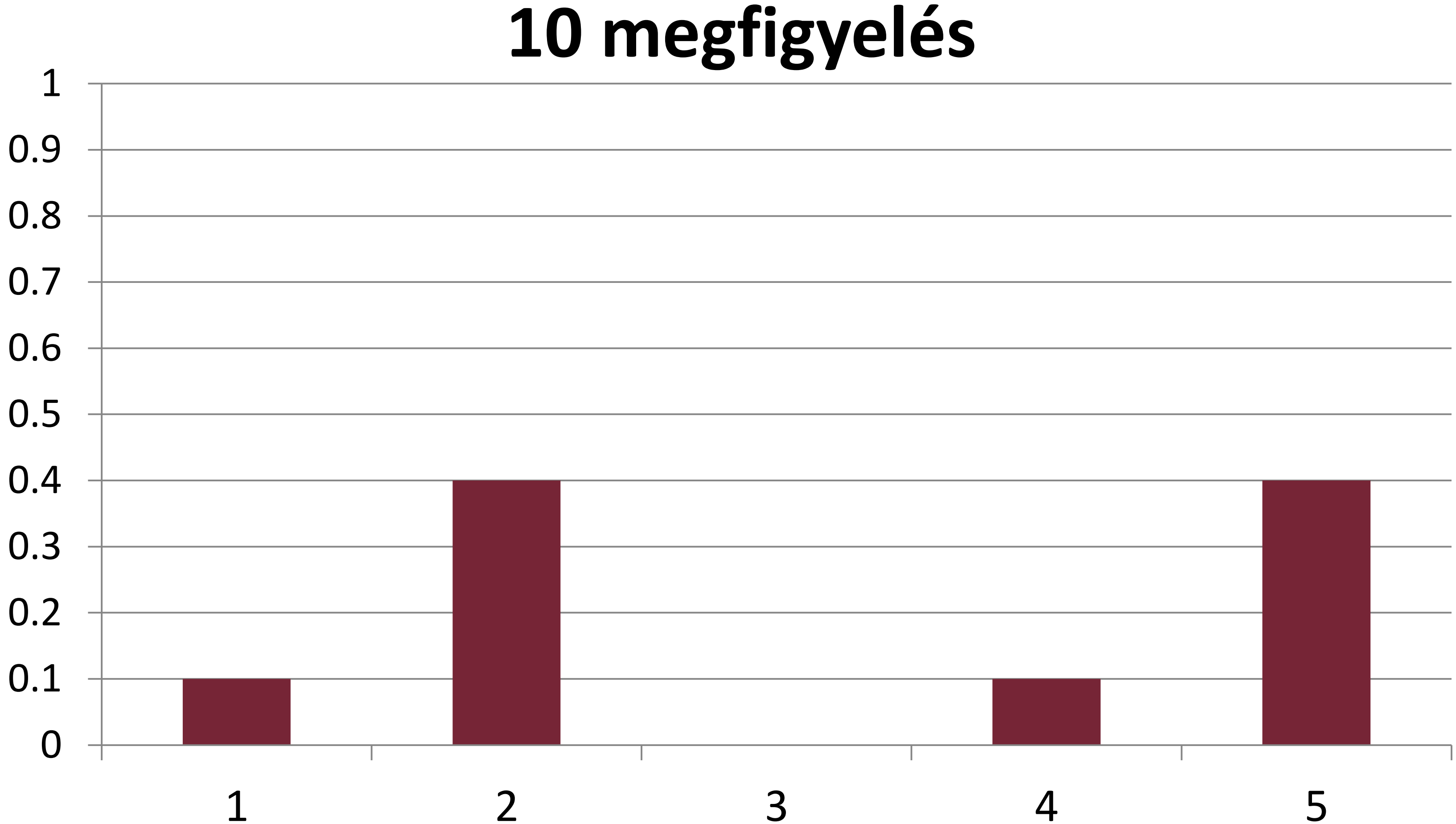
- Mérjük meg egy jellemző értékét! (1..5)



Számított átlag: 3,6

Számszerű jellemzők mérése

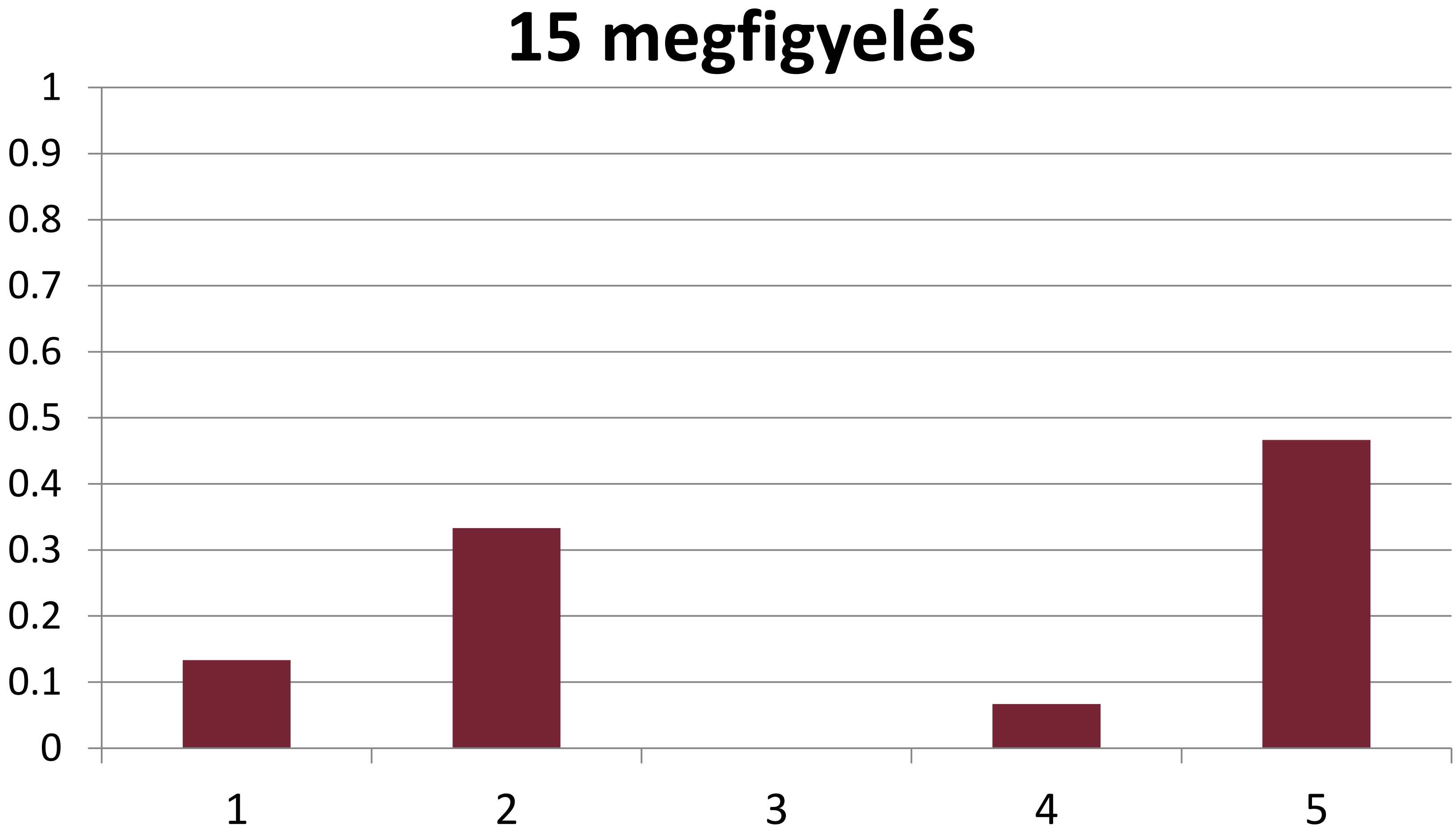
- Mérjük meg egy jellemző értékét! (1..5)



Számított átlag: 3,3

Számszerű jellemzők mérése

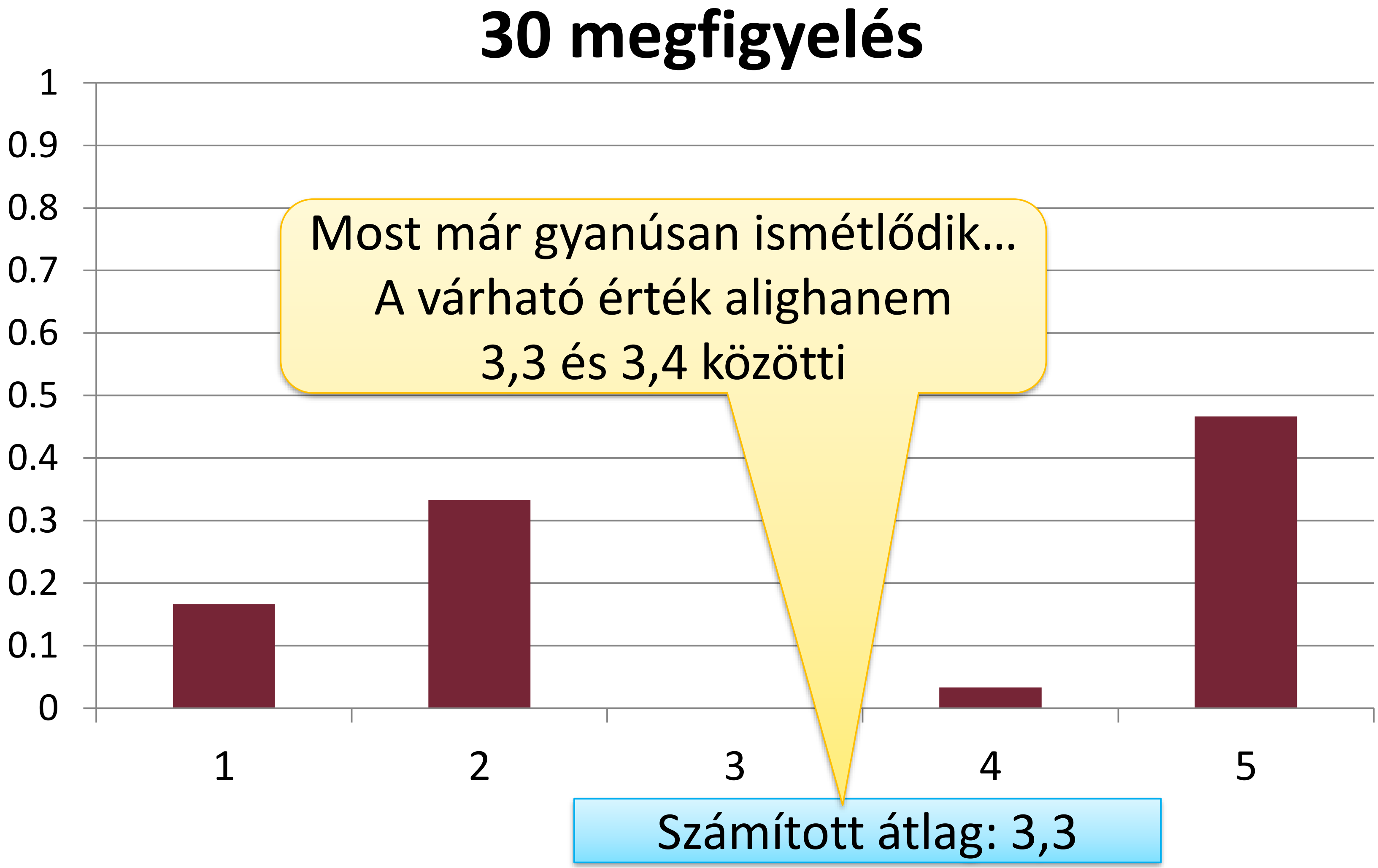
- Mérjük meg egy jellemző értékét! (1..5)



Számított átlag: 3,4

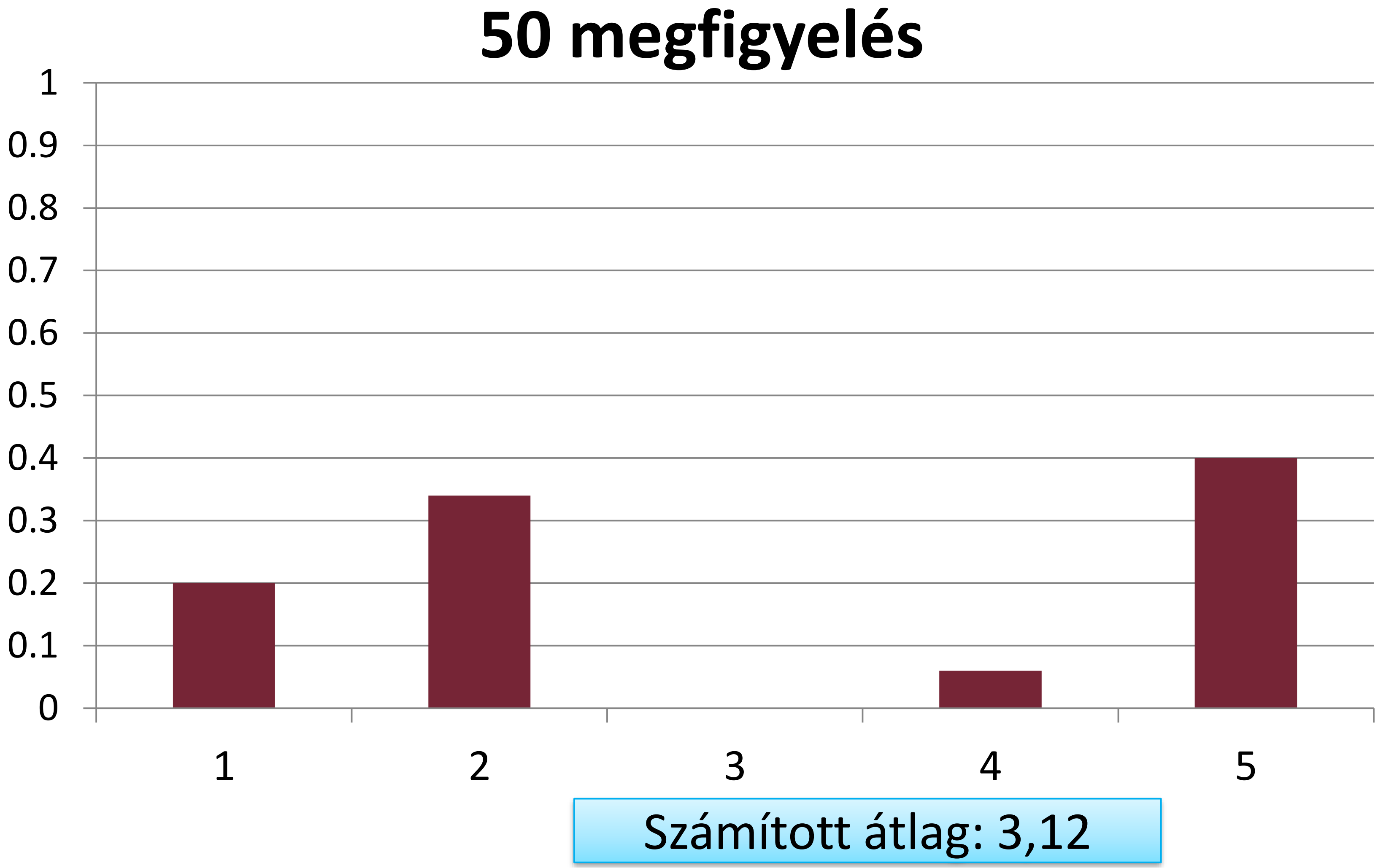
Számszerű jellemzők mérése

- Mérjük meg egy jellemző értékét! (1..5)



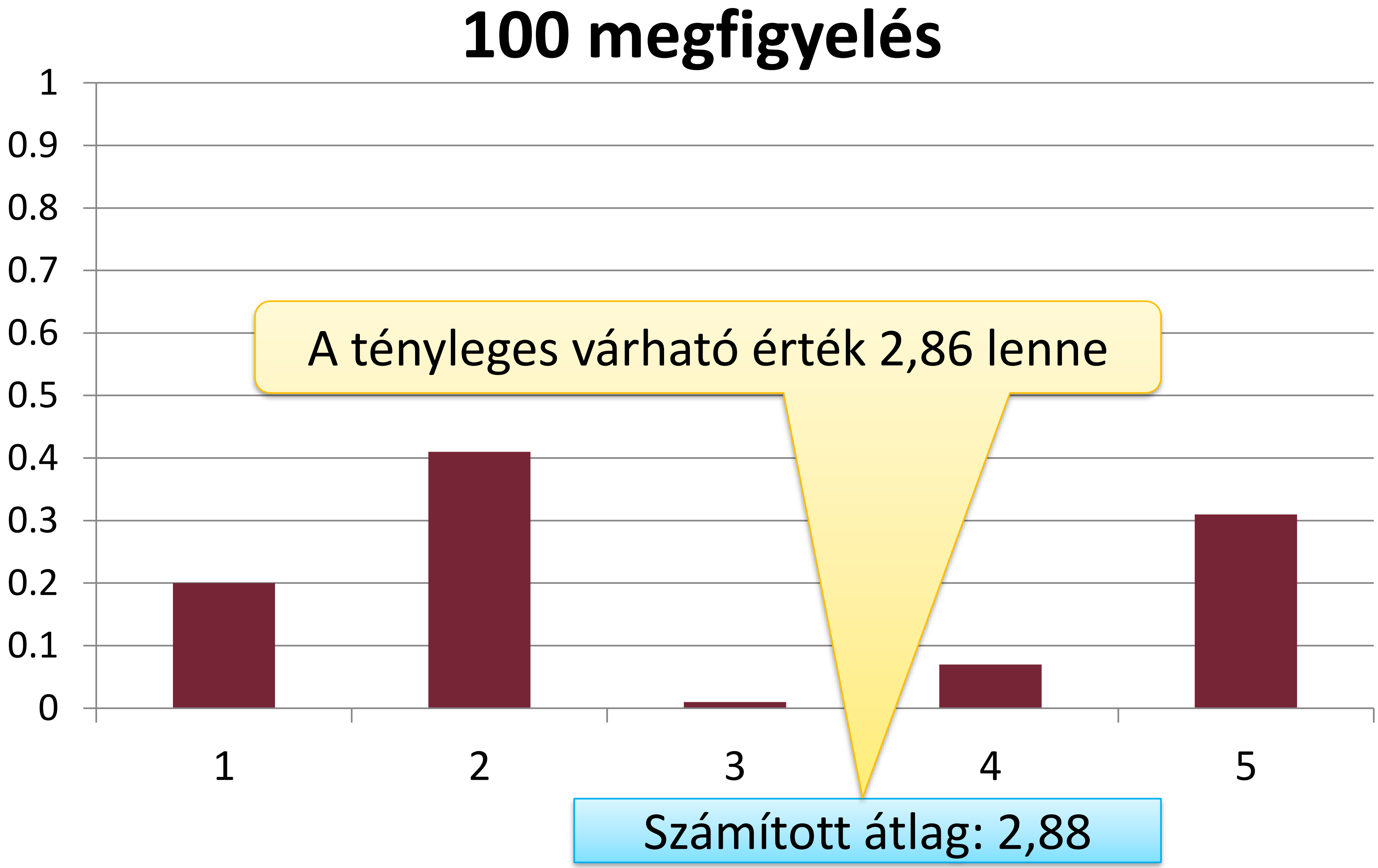
Számszerű jellemzők mérése

- Mérjük meg egy jellemző értékét! (1..5)



Számszerű jellemzők mérése

- Mérjük meg egy jellemző értékét! (1..5)



Ismétlés: tapasztalati átlag, szórás

- Valószínűségi változó: E (vizsgálandó jelenség)
  - Várható érték: $\mu = \mathrm{E}(X)$ átlagos viselkedés)
  - Szórás: $\sigma = \sqrt{\mathrm{E}(X-\mu)^2}$ (eltérések mértéke)
- Mintavétel: x_1, x_2, … x_t (mérések, megfigyelések)
  - Tapasztalati átlag: $\bar{x} = \dfrac{x_1 + x_2 + \ldots x_t}{t}$
  - Szórásra **nem jó**: $\sqrt{\dfrac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\ldots(x_t-\bar{x})^2}{t}}$
  - Korrigált tapasztalati szórás: $\sqrt{\dfrac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\ldots(x_t-\bar{x})^2}{t-1}}$

Tapasztalati átlag

- Módszer: számtani átlag képzése
  - Ismételt megfigyelések
  - Egymástól függetlenül
  - Azonos feltételek mellett
- Kérdések
  - Hány megfigyelést kell végezni?
  - A tapasztalati átlag mennyire jellemzi a valódi várható értéket?
- Először tisztázandó:
  - *A tapasztalati átlag eloszlása*

Tapasztalati átlag eloszlása

- Kísérlet = megfigyelések sorozata
- Megfigyelések sorozatának **tapasztalati átlaga**:
  - Egy jellemzőt ***t*** db független megfigyeléssel mérve,
  - majd a mért értékeket átlagolva kapott eredmény
- **Centrális határeloszlás tételé**ből következik:
  - Tetszőleges eloszlású jellemző
  - (de legyen *véges* ***m*** várható értékű és **s** szórású)
  - tapasztalati átlaga **t**→∞ esetén közelítőleg
  - **normális eloszlású**,
  - μ = ***m*** várható értékkel és $\sigma = {}^{\mathbf{s}}/_{\sqrt{t}}$ szórással

Tapasztalati átlag eloszlása

- Kísérlet = megfigyelések sorozata
- Megfigyelések sorozatának **tapasztalati átlaga**:
  - Egy jellemzőt ***t*** db független megfigyeléssel mérve,
  - ...tt eredmény
- **C**... következik:
  - ...
  - (de legyen *véges* ***m*** várható ... ékű és **s** szórású)
  - tapasztalati átlaga **t**→∞ esetén közelítőleg
  - **normális eloszlású**,
  - μ = ***m*** várható értékkel és $\sigma = {}^{s}/_{\sqrt{t}}$ szórással

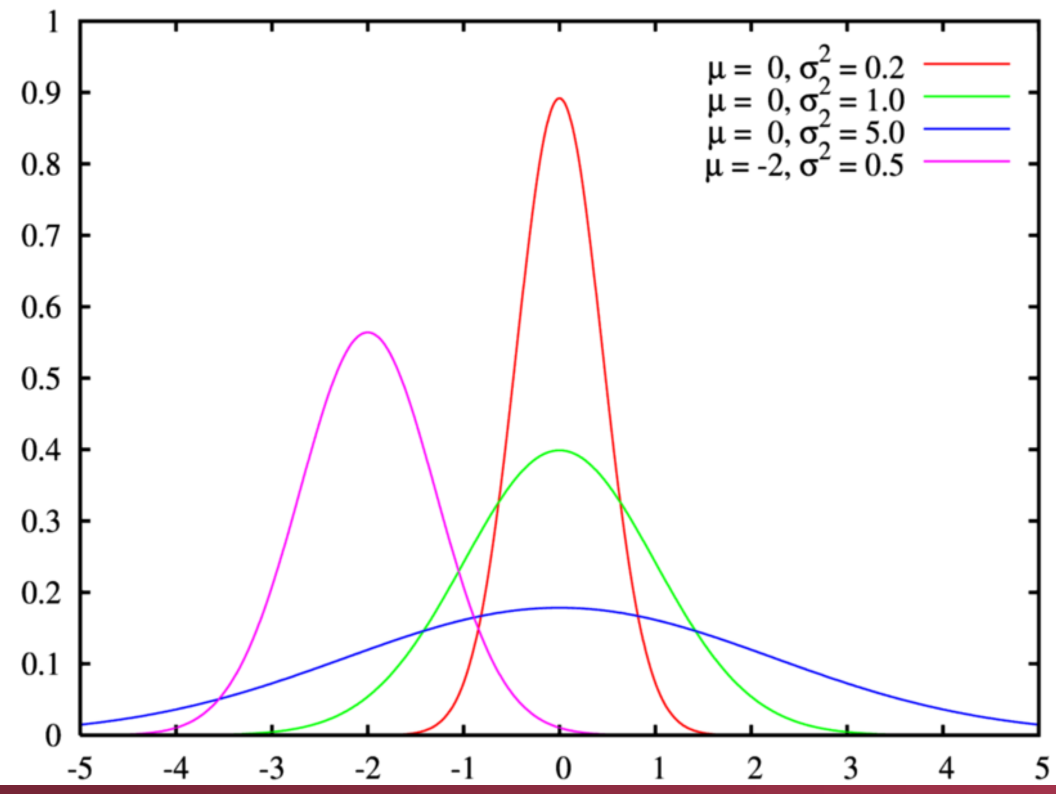
Ökölszabály:
ismert szórásnál t>30,
ismeretlen szórásnál t>100
után kezd elfogadható lenni a közelítés

A normális (Gauss) eloszlás

- Valószínűségsűrűség-függvénye: (nem kérdezzük)

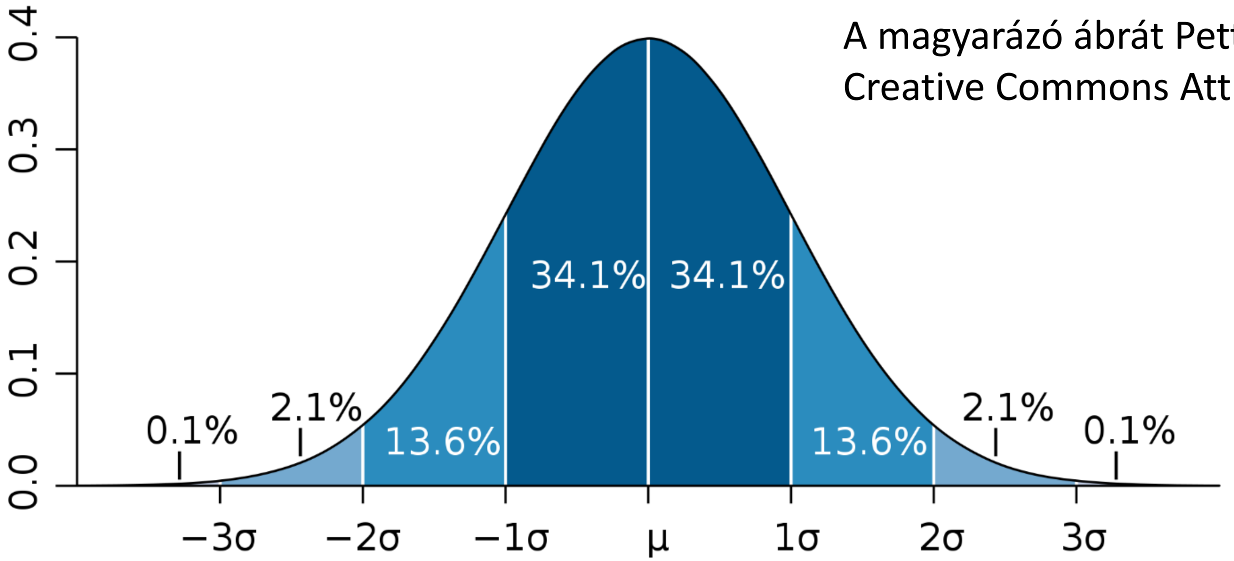
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- Paraméterek
  - Várható értéke μ
    - → **m** a mi esetünkben
  - Szórása σ
    - → $\mathbf{s}/\sqrt{\mathbf{t}}$ esetünkben



A normális (Gauss) eloszlás

- A várható érték körül koncentrálódik



- A normális eloszlású változó…
  - az esetek **68%**-ában legfeljebb **1σ** messze kerül μ-től
  - az esetek **95%**-ában legfeljebb **2σ** messze kerül μ-től
  - az esetek **99,7%**-ában legfeljebb **3σ** messze kerül μ-től
  - …

Centrális határeloszlás tétele

- CLT (Central Limit Theorem)
  - A minták statisztikáinak átlaga normális eloszlást követ (bizonyos feltételek mellett).
  - Bővebben (és alaposabban): **Valószínűségszámítás**
  - $\bar{x} \sim N\left(mean = \mu, \sigma = \frac{s}{\sqrt{n}}\right)$
    - $\bar{x}$ a mintaátlag
    - μ a populáció várható értéke
    - s a populáció (empirikus) szórása
    - n a mintaméret

Konfidenciaintervallumok

- Ha tehát a tetszőleges eloszlású, ***s*** szórású vizsgált jellemzőről ***t*** db (>30) megfigyelést végzünk
- A tapasztalati átlagáról…
  - **68%** biztonsággal kijelenthető, hogy legfeljebb $\boldsymbol{s}/\sqrt{\boldsymbol{t}}$ pontatlansággal becsli ***m*** értékét
  - **95%** biztonsággal $\boldsymbol{2s}/\sqrt{\boldsymbol{t}}$ sugarú intervallumba esik
  - **99,7%** biztonsággal $\boldsymbol{3s}/\sqrt{\boldsymbol{t}}$ sugarú intervallumba esik
- És ***t*** növelésével gyökösen szűkül az intervallum

Konfidenciaintervallumok

- Ha tehát a tetszőleges eloszlású, ***s*** szórású vizsgált jellemzőről ***t*** db (>30) megfigyelés[…]égzünk
- A ta[…]gáról…
  - **68%** biztonsággal kijele[…] legfeljebb $s/\sqrt{t}$ pontatl[…]
  - **95%** biztonsággal $2s/\sqrt{t}$ sugarú intervallumba esik
  - **99,7%** biztonsággal $3s/\sqrt{t}$ sugarú intervallumba esik
- És ***t*** növelésével gyökösen szűkül az intervallum

Konfidenciaszint

Egyedi megfigyelés szórása

Konfidenciaintervallum sugara (félszélessége)

Kísérlettervezés példa

- A várható értékre 30 megfigyelés
  - Tapasztalati átlag: 2,3 s (jó-e ez? kell még mérni?)
  - Tapasztalati szórás: $\boldsymbol{s}$ = 1,1 s
- Cél
  - 99,7%-os konfidenciaintervallum 0,6 s széles legyen
- Kísérlettervezés
  - Elvárt sugár (félszélesség) = 3σ = $\boldsymbol{3s}/_{\sqrt{\boldsymbol{t}}}$ < 0,3 s
    - (ez a σ az átlag szórása, nem az eredeti mért jellemzőé!)
  - Ezért t = 121 megfigyelés kell legalább
- Hol a csalás?

Korrekció

- Többnyire a tényleges eloszlás paraméterei *a priori* ismeretlenek (különben minek mérnénk?)
- Így nem használható fel a tényleges ***s*** szórás
- Csak a tapasztalati szórás használható → Gauss/normális helyett Student t-eloszlás
  - (más konfidenciaintervallumok)
- $t \to \infty$ esetén Student → normális
- Ökölszabály: t>100 esetén használható a Gauss