

6. gyakorlat – Követelmények elemzése, felderítő adatelemzés

1. Szerver teljesítménye – felderítő adatelemzése

Egy szerveren az alábbi teljesítményjellemzőket mértük:

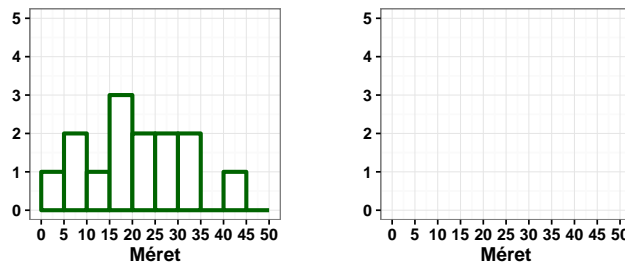
Mintavétel időpontja [ms]	500	600	700	800	900
Utolsó 100ms alatt feldolgozott kérések száma [darab]	11	12	21	18	20
Utolsó 100ms átlagos kiszolgálási ideje [ms]	15	20	21	25	27
Utolsó 100ms CPU kihasználtság [%]	12	13	16	17	19
Utolsó 100ms HDD I/O kihasználtság [%]	55	63	87	61	73

- Ábrázoljuk a feldolgozott kérések számát és a CPU kihasználtságot pontfelhő (scatterplot) diagramon! Értelmezzük a diagramot!
- Az első mintavétel idején mekkora az átbocsátási ráta értéke? Az 5 mintavétel alapján mekkora az átbocsátási ráta tapasztalati átlaga és mediánja? Mi tartozik a 40%-os kvantilisbe?
- Vajon mely mért jellemzők között sejthető ok-okozati viszony?

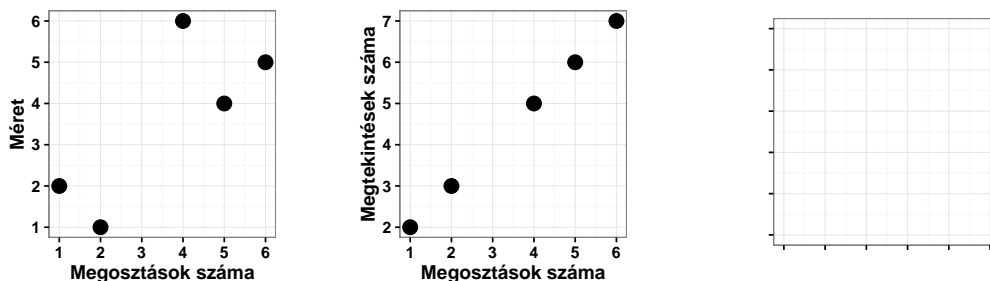
2. Képgaléria – adatelemzés

Online képgalériánkban a felhasználók keresés alapján megjeleníthetnek a keresőkifejezésre illeszkedő képeket.

- Az alábbi hisztogramon ábrázoltuk az albumok méretének eloszlását. Mivel a tárhely hatékony szervezéséhez elég azt tudnunk, hogy hány 10 alatti, 10 és 20 közötti stb. képet tartalmazó albumunk van, az alábbihoz képest kétszeres oszlopszélességű hisztogramot szeretnénk (szintén a 0 mérettől kezdve felszámítva az oszlopokat). Rajzoljuk meg az ábrát!



- Pont-pont diagramon (scatterploton) ábrázoltuk 5 kiválasztott album méretét illetve megtekintési számát a megosztási számmal összehasonlításban. Igaz-e, hogy minél nagyobb az album, annál többen tekintik meg? Válaszolja meg a kérdést egy harmadik pont-pont diagramon, amely a megtekintések számát a méret függvényében ábrázolja!



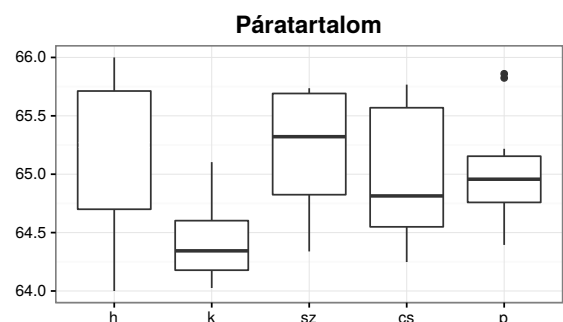
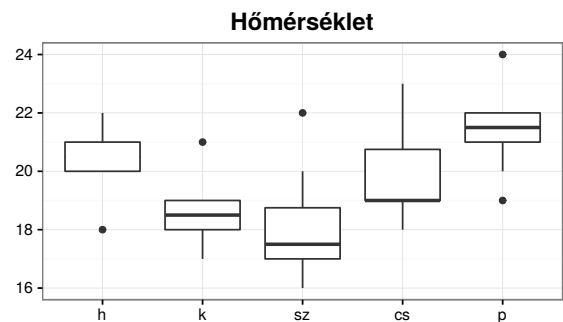
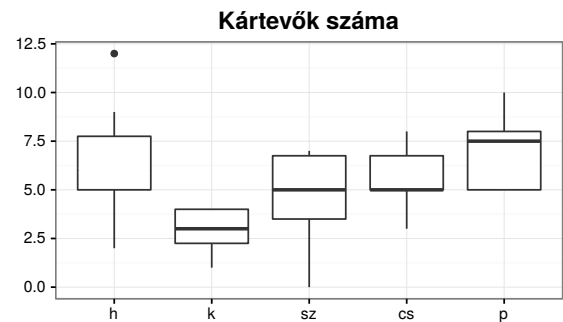
- Az albumok jellemző népszerűségét szeretnénk meghatározni, emiatt a pont-pont diagram alapján kiszámoltuk a megtekintési számok átlagát és mediánját. Általánosságban megtehető-e ez egy pont-pont diagram alapján? Mennyivel változnak ezen középértékek, ha feltöltünk egy új albumot, amelyet 40-en tekintenek meg?

3. Szenzorhálózat (korábbi zh feladat) – adatelemzés

Adott egy mezőgazdasági szenzorhálózat, amellyel a szabadföldes, üvegházi, ill. fóliasátras területeink állapotát követjük nyomon a mért értékek (hőmérséklet, páratartalom, fényerősség, szélsébség, detektált kártevők stb.) alapján.

Dátum	Hőm. [°C]	Pára. [%]	Kártevők [db]
2015. 05. 04. 08:00	18	66,00	3
2015. 05. 04. 09:00	20	65,75	6
2015. 05. 04. 10:00	20	65,75	8
2015. 05. 04. 11:00	20	65,50	9
2015. 05. 04. 12:00	20	65,50	5
2015. 05. 04. 13:00	21	65,00	12
2015. 05. 04. 14:00	21	64,70	5
2015. 05. 04. 15:00	21	64,70	6
2015. 05. 04. 16:00	21	64,60	7
2015. 05. 04. 17:00	22	64,00	2

- Sajnos a május 4. hétfői középértékek (medián) lemaradtak az ábráról, rajzoljuk őket be a táblázatban található adatok alapján!
- Értelmezze a diagramokat: mely változó(k) első kvartilisei mutat(nak) szigorúan monoton változást az idő folyamán?
- (Kiegészítő feladat.) Szeretnénk párhuzamos koordináta diagramon összevetni a hétfői hőmérsékleti értékeket a detektált kártevők számával.



4. Szenzorhálózat (korábbi zh feladat) – teljesítményelemzés (*)

(A 3. feladathoz kapcsolódó teljesítményelemzési feladatok.) A különböző típusú szenzorok a helyüktől számított 100 méteres körzetben lévő területekről szolgáltatnak adatokat. A szenzorok mérési eredményeiket időbélyeggel ellátva, rádiós kommunikációs hálózaton továbbítják a központnak. A központi számítógép processzora feldolgozza a kéréseket, majd archiválási cézzal kiírja őket egy tárolóegységre. A gazdaságunk összesen 4500 szenzort telepített, amelyek percnként egy-egy mérési eredményről adnak jelentést. A rendszer sikerrel kiszolgálja a terhelést. A rádiós kommunikációs hálózat 100 mérési eredményt képes másodpercnként továbbítani. A központi számítógép CPU idejének 75%-a tétlenül múlik. A tárolóegységet 8 ms-ig foglalja le minden egyes kérés kiírása.

- Másodpercnként hány mérési adat a rendszer jelenlegi átbecsátása?
- Mekkora a hálózat, CPU, ill. tároló átbecsátása, átbecsátóképessége és kihasználtsága?
- A mérési pontosság javításához hány szenzort helyezhetünk még üzembe ugyanezen a területen az infrastruktúra fejlesztése nélkül? Feltételezzünk lineáris skálázódást!
- A rádióhálózat ügyes kódolással biztosítja, hogy egyszerre több szenzor is sugározhasson mérési eredményeket. Átlagosan hány szenzor rádiója sugároz egyszerre (vagyis hánszoros az átlapolódás) a hálózaton jelenleg, ill. a hálózat maximális terheltsége esetén, ha egy mérési eredmény sugárzása 40 ms-ig tart?

5. Vasúti biztosítóberendezés követelményelemzése

Vasúti biztosítóberendezést tervezünk. A rendszer elsődleges célja a vonatok összeütközésének megakadályozása. A megfelelő rendszer kifejlesztésének kulcsa a jó minőségű követelményspecifikáció, ugyanis a követelmények alapján kell majd teszteseteket és egyéb ellenőrző vizsgálatokat kidolgoznunk.

1. táblázat. A vasúti biztosítórendszer követelményei (részlet)

R1	Biztonság	A felügyelt pályarendszeren tartózkodó vonatok nem ütközhetnek össze.
R2	Működés	A vonatoknak biztosítani kell, hogy elérhessék az úti-céljukat.
R3	Optimalitás	Minimalizálni kell a vonatok menetidejét.
R4	Pályaszakaszok felügyelete	A pályarendszert szakaszokra kell osztani, ezeken egyszerre egy vonat tartózkodhat.
R5	Szakaszokra bontás	A pályarendszert szakaszokra kell bontani.
R6	Foglaltság	Egy szakaszon egyszerre egy vonat tartózkodhat.
R7	Foglaltság érzékelése	Valamilyen módon érzékelni kell, hogy egy szakaszon áll-e vonat, vagy nem.
R8	Hibatűrés	A komponensek meghibásodására fel kell készülni.
R9	Foglaltságjelző szenzorok	A foglaltságot többféle, redundánsan kialakított szenzorral kell érzékelni.
R10	Sínekbe épített szenzor	A sínekbe mindegyik szakaszon szenzorokat kell telepíteni, amik jelzik, hogy a szakaszon áll-e vonat, vagy sem.
R11	Kamerás rendszer	Ahol lehetséges, kamerákat kell telepíteni a szakaszok megfigyelésére.
R12	Helyzetmeghatározás	A vonatoknak folyamatosan jelezni kell a helyzetüket a központi vezérlő felé.
R13	GPS alrendszer	A vonatokat GPS alrendszerrel kell felszerelni.
R14	Vezeték nélküli kapcsolat	Biztosítani kell, hogy a vonatok vezeték nélküli hálózaton jelezhessék a helyzetüket a központi vezérlő felé.
R15	Vonatok vezérlése	Meg kell tudni akadályozni, hogy a foglalt szakaszra másik vonat is ráhajthasson.
R16	Vonat leállítása	A központi rendszer azonnal leállíthatja a vonatot.
R17	Mozdonytípusok támogatása	A rendszernek támogatnia kell minden, a sínpáron közlekedni képes mozdonytípust.
R18	Mozdony nem módosítható	Nem alkalmazható olyan megoldás, amihez a mozdonyok vezérlését meg kellene változtatni.

- Gyűjtsük össze azokat a szereplőket, akik egy ilyen rendszer kifejlesztése kapcsán érintettek, vagyis követelményeket támaszthatnak a leendő rendszerrel szemben (ún. *stakeholderek*)!
- A stakeholderek felderítése után összegyűjtöttük az általuk támasztott követelményeket is, ennek egy részletét tartalmazza az 1. táblázat. Rajzoljuk fel a követelmények közötti függőségi viszonyokat egy gráf segítségével! A gráfban A -ból B -be mutató irányított éllel jelezzük, ha (1) az A követelmény a B követelmény része (*kompozíció*), (2) az A követelmény finomítja (részletezi) a B követelményt (*refines* kapcsolat), illetve (3) az A követelmény származtatható a B követelményből (*derive* kapcsolat). Ne foglalkozzunk azzal, hogy két követelmény közül ezen viszonyok melyike áll fent; most csak a kapcsolat megléte a fontos.
- A felsoroltak közül melyek funkcionális követelmények, illetve milyen típusúak az extrafunkcionális követelmények (biztonságosság, teljesítmény, megbízhatóság stb.)?
- Vizsgáljuk meg, hogy konzisztens-e a bemutatott követelményrendszer! Ha nem az, akkor mutassunk példát ellentmondásra!
- A fentiekből adjunk példát közvetlenül ellenőrizhető követelményre!

Kiegészítő feladatok

6. Közösségi oldal

Internetes közösségi oldalt működtetünk. Az utóbbi időben számottevően népszerűbb lett az oldal, de ezáltal a válaszidő is kellemetlenül megnőtt. Az üzleti cél, hogy csúcsidőszakban egyszerre 1500 felhasználót átlagosan négy másodperces válaszidővel szolgáljon ki a honlap.

- a) Minimálisan mekkorára kell tervezni a kiszolgáló infrastruktúra átbecsülőképességét, ha az azon kívüli késleltetés (hálózati forgalom, HTML megjelenítés a kliensoldalon) egy másodpercnél becsülhető?
- b) Az újratervezett weboldalon a mérések szerint egyetlen kérés kiszolgálása átlagosan 20 ms CPU-időt igényel a webszerveren, és 12,5 ms erejéig foglal le egy adatbázisszervert. Jelenleg 15 webszerver fogadja a kéréseket és az adatbázis 5 kiszolgálóra van replikálva. Lineáris skálázhatóságot feltételezve, milyen számítógépből és mennyit kell még legalább venni a fenti cél eléréséhez?
- c) (*) A kibővített rendszerben mekkora lesz az egyes szervertípusok kihasználtsági aránya? Ha az a cél, hogy még a csúcsidőszakban is legfeljebb 50%-os legyen a kihasználtság, meddig kellene még bővíteni a rendszert?
- d) Tekintsünk csak 2 db webszervert és 3 db adatbázis szervert. Készítsünk állapot alapú modell(ek)e(t), amely(ek) az infrastruktúra erőforrásait modellezi(k) az elérhetőségeik (szabad/foglalt) szerint. Milyen tervezői döntésekkel szembesülünk? Mik az egyes lehetőségek előnyei és hátrányai?