

Vágjunk utat az adatdsungelben!

Mekkora egy adattábla?

Gondoljunk csak bele, hogy az előző példák közül mondjuk a vasúti vagy a repülőtéri adatbázisnak mennyi járatról kell adatokat tárolni, és járatonként milyen sokat! Érezhető, hogy a járatadatokat tároló tábláknak nagyon sok sorból és oszlopból kell állniuk. Amikor még nem terjedtek el az internetes adatbázisok, vagyis néhány évtizede, a magyar vasúti mentrendet évente több száz oldalas könyvként nyomtatták ki (egészen 2017-ig létezett nyomtatott változat). Ha ebben kellene a Szolnok–Békéscsaba viszonylat kívánt adatait (IC, másodosztály, holnap) megtalálni, az bizony rengeteg keresgéléssel járna, és ezért hosszú ideig tartana.

Persze lehetne azt mondani, hogy töröljük ki minden más adatot a táblából, és akkor sokkal könnyebb dolgunk lesz. Ez méltánylandó ötlet, bár a táblát összeállítók dühöngnének, hogy hiába dolgoztak az adatbevitellel annyit. Ha ez a kifogás sem tartana vissza az adattörléstől, akkor gondoljunk a következőre! Mihez kezdenénk, ha a mai lecke végén egy másik járat, mondjuk a Székesfehérvár–Veszprém adatairól érdeklődné egy feladat?

Alkalmazzunk szűrőt!

Amikor szüretkor préselik a szőlőt, a kicsorgó gyümölcslé magával sodorhat némi szőlőhéjat és -magot, esetleg a fűtök közé került szennyeződéseket, rovarokat. Ahhoz, hogy ezektől ne romolhasson meg a lé, át szokták szűrni, hogy a szűrőn fennakadjon a felesleg. Sok más folyamatban is alkalmazunk szűrőket a kellő és a fölösleges dolgok elkülönítésére, például bizonyára már mindenki látta, miként szűrik le a megfőtt tésztát. A tészta nem fér át a szűrő lyukain, viszont a főzővíz azokon át el tud távozni.



Szinte hallom a reklamációkat, hogy ez két különböző dolog, hiszen a gyümölcslénél az a szemét, ami fennmarad a szűrőn, a tésztánál azonban a fennmaradó részből lesz sajtos vagy mákos tészta, esetleg spagetti. A méltatlankodásnak van némi alapja, de ha úgy nézzük, hogy szétválasztjuk a szükséges és szükségtelen dolgokat, akkor mindegy, melyik marad fenn a szűrőn, a lényeg a szétválasztás.

A digitális világban az utóbbi módszer működik, az **adatbázisok szűrői** azokat a sorokat mutatják meg, amelyek a szűrési feltételnek vagy feltételeknek megfelelnek, a többit elrejtik a szemünk elől. Hogy a dolog kellően szemléletes legyen, gondoljunk a fejezet első példájára: az iskola tanulóinak adattáblájából kiszűrhetjük a 10.g tanulóit, így máris lesz egy osztálynévsorunk. Itt álljunk meg egy pillanatra! Nem szeretnénk elvarratlan szálakat hagyni.

Rendezés

A 10.g-ben tanulók nevének listája nem lesz automatikusan a hagyományos értelemben vett osztálynévsor, hiszen ahhoz a neveket ábécérendben kellene felsorolni. Az összes adat közül kiválogatottnál szükség lehet többféle sorrendre.

Például ilyen sorrend lehet:

- neveknél az ábécérend;
- tornasornál a testmagasság;
- orvosi váróban a betegek érkezési időpontja;
- online áruházban a termékek ára;
- lakberendezési áruházban a kész függönyök magassága;
- vonatjáratoknál az indulási idő;
- internetes keresőprogramnál a keresésnek megfelelő oldalakat felkeresők száma;
- településeknél az ott élők lélekszáma.



Az adatbázis-kezelők persze a **rendezésre** is lehetőséget biztosítanak, az online felületeken általában csak konkrét rendezési módok között választhatunk. Ez általában rendben is van így, hiszen ki is szeretné a **találatokat** (a szűrő által kiválogatott rekordokat) mondjuk az árut beszállító életkora vagy az áru és a csomagolás súlyaránya szerint tanulmányozni. De ezek a lehetőségek is arról tanúskodnak, hogy az adatokat lehet rendezni.

Irányított szűrés

Ma már szinte elképzelhetetlen lenne az életünk az online adatbázisok nélkül, hiszen a mindennapjainkban meglehetősen sűrűn fordulunk hozzájuk. A teljesség igénye nélkül lássunk egy elég tág felsorolást! Online adatbázishoz fordulunk, ha

- egy település vagy egy utca helyét keressük a térképen;
- vásárolnánk egy online áruházban;
- szeretnénk megtudni, hogy holnap mikor indul távolsági busz vagy vonat Egerbe;
- szeretnénk mozi- vagy koncertjegyet venni;
- egynémely hivatalos ügyünket szeretnénk elintézni, mondjuk letiltani az elveszett bankkártyánkat;
- képet keresünk a neten;
- klipet keresünk kedvenc videómegosztókon;
- egy új applikációt szeretnénk feltelepíteni a tabletünkre vagy a mobilunkra;
- ebédet, vacsorát, netán pizzát rendelnénk valahonnan;
- szeretnénk megnézni, hogy milyen jegyeink vannak földrajzból.

Az eddig leírtak azt sugallják – és ha belegondolunk, így is van –, hogy ezekben az adatbázisokban tényleg irdatlan mennyiségű adatot tárolnak. Ha ezeket egyszerre a nyakunkba zúdítanák, szinte lehetetlen lenne kiválasztanunk belőle a számunkra szükséges adatokat. Valami olyasmi lenne, mintha a testvérünk a születésnapja előtt a kezünkbe nyomná a *Háború és béke* meglehetősen vaskos kötetét azzal a megjegyzéssel, hogy aláhúzott benne egy szót, ami arra utal, hogy mit is szeretne ajándéka.

Amikor egy ilyen adatbázist használunk, nem mindig vagyunk tudatában annak, hogy egy nagy táblázatban keresgélünk. Ennek persze az az oka, hogy az oldalon nem mutatják a táblázatrácsot, hisz az adatok elhelyezkedése már tájékoztat, így az emiatt fölöslegessé vált rácsot elrejtik.

Ebben a kásahegyben szinte esélyünk se lenne rátalálni a minket érdeklő adatokra, ha az oldalak üzemeltetői nem kínálnának fel szűrőket. Például a termékcsoportokat a webáruházakban vagy az utazási viszonylatokat a menetrendekben. Ezek persze a tervezőik elképzelései szerint szabályozottak, nem kérhetünk bármilyen szűrést. Néha ez bosszantó, de általában kellő változatosságot biztosítanak, hogy megtalálhassuk a kívánt adatokat. Aszúrt adatok mindig valamiféle rendezettséggel jelennek meg a képernyőn, ez lehet például ár szerinti a webáruházakban, indulási idő a menetrendekben stb.

Még a szűrés ellenére is előfordulhat, hogy rengeteg adat közül lennénk kénytelenek válogatni, gondoljuk csak el, hogy például egy elektronikai webáruház kínálatában több száz fajta monitor szerepel. Ilyenkor általában lehetőségünk van további segédszűrők alkalmazására. Például ilyen lehet:

- értékhatár;
- márka;
- kategória (pl.: tejtermékek → tejek, sajtok...);
- méret;
- szín;
- ügcsoport;
- a keresett zene előadója;
- a keresett kép mérete;
- átszállás nélküli járat a vonatoknál.

Néha arra is van lehetőség, hogy a rendezés szempontját átállítsuk, mondjuk a monitor ára szerint a monitor képátlójának mérete szerintre. Ezek is gyorsíthatják a keresett adatok közül a számunkra megfelelő kiválasztását.

Ezek a lehetőségek többnyire elégségesek, de néha bosszantó, hogy pont olyan szűrést nem lehet kérni, ami nekünk a legmegfelelőbb lenne. Például a vasúti menetrendben nem lehet csak a délutáni vonatokra keresni. De ennyi kényelmetlenségért biztos kárpótol minket az a kényelem, hogy rengeteg dolgot otthonról intézhetünk. Bizony néhány évtizede még telefonon kellett érdeklődni a vonatok indulásáról, esetleg elzarándokolni a vasútállomásra és ott informálódni, ha nem vettük meg otthonra az adott évi vasúti menetrendet.



Feladatok

1. Keressünk fel két-három elektronikai áruházat, UHD (Ultra High Definition – nagyfelbontású) televíziót vásárolnánk maximum 110 cm-es képátlóval! Írjuk fel áruházanként a kívánt készülékek legmagasabb és legalacsonyabb árát! Jegyezzük fel azt is, hogy hol milyen segédszűrés, rendezés lehetséges!
2. Keressünk képet a Google képkeresőjével, mégpedig légi felvételt a Margit-szigetről. Állítsuk be a következő segédszűrőket: a kép legyen nagyobb 2 MP-nél és álló formátumú! A találatok közül egyet mentsünk le a gépünkre!
3. Látogassunk el egy online élelmiszer-áruházba! Jegyezzük fel, hogy milyen termékkategóriák vannak, nézzük meg és írjuk fel az egyik termékcsoport kategóriáit is!