

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR

Szerzők:

Kovács József – Hatvani István Gábor – Magyar Norbert
Tanos Péter – Trásy Balázs – Borbás Edit
Garamhegyi Tamás – Vid Gábor – Kovácsné Székely Ilona

FÖLD- ÉS KÖRNYEZETTUDOMÁNYI SZÁMÍTÁSOK

I.

Szerkesztette:

KOVÁCS JÓZSEF



**EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR**



**FÖLD- ÉS KÖRNYEZETTUDOMÁNYI
SZÁMÍTÁSOK I.**

Szerzők:

Kovács József, Ph.D.

Hatvani István Gábor, Ph.D.

Magyar Norbert

Tanos Péter

Trásy Balázs

Borbás Edit

Garamhegyi Tamás

Vid Gábor

Kovácsné Székely Ilona, Ph.D.

Budapest,

2015

Szerkesztette:

Dr. Kovács József

Lektorálták:

Szakmai lektor: dr. Zibolen Endre és Zibolen Erzsébet

Nyelvi lektor: Tanos Márton

Kiadja: az Eötvös Loránd Tudományegyetem – Természettudományi Kar

ISBN: 978-963-284-676-7

Felelős kiadó: Dr. Surján Péter dékán, ELTE TTK

Terjedelem: 76 oldal

Tartalom

Előszó	3
1. Hasznos billentyűkombinációk és cellahivatkozások	5
2. Feltételes függvények alkalmazása felszínalatti idősorokon	12
3. Szöveg- és dátumkezelő függvények	18
4. Adatrendeztetés és importálás, egy informatikai példa	26
5. Az FKERES függvény használata felszíni vízminőségi adatok esetén	33
6. Az FKERES és VKERES függvény alkalmazása harmatpont meghatározására, meteorológiai adatokból	37
7. Részösszegképzés a troposzférikus ózonkoncentráció-adatokkal	43
8. Keresztkorreláció és autokorreláció	52
a. Keresztkorreláció	52
b. Autokorreláció	56
9. Célértékkeresés	62
10. Összefoglaló feladat – a súlyozás problémaköre	67
Köszönetnyilvánítás	73
Felhasznált irodalom	74

Előszó

Az elmúlt évtizedekben a környezeti rendszerek tulajdonságainak részletes, némely esetben automatizált mérések segítségével történő megismerése a természettudományok számos területén előtérbe került. A megfelelő vizsgálatokból származó minták terepi és laboratóriumi mérései nagy elemszámú adathalmazokhoz vezetnek. Létrehozásukat azonban hazánkban nem feltétlenül követi – közvetlenül – kellő mélységű kiértékelés, így Magyarország adatokban gazdag, de információban szegény országnak tekinthető (HATVANI ET AL., 2014). Hisszük és reméljük, hogy a közeljövőben ez a trend megváltozik és az adatelemzés a természettudományokban (és más szakterületeken is) egyre jelentősebb szerephez jut. Meggyőződésünk, hogy az adathalmazok rendezésének és kezelésének ismerete minden természettudományi (ill. gazdasági, mérnöki stb.) végzettségű szakember számára elengedhetetlen.

Az Eötvös Loránd Tudományegyetem Természettudományi Karának több szakán is hosszú évek óta folyik az alapvető adatelemző ismeretek oktatása mind B.Sc., mind pedig M.Sc. hallgatók számára. A hallgatók graduális képzésük során átfogó elméleti oktatásban részesülnek, azonban az itt megszerzett tudás nem párosul minden esetben megfelelő mélységű gyakorlati ismeretekkel. Megismerik a különböző szférák mintavételezésének részleteit, de a létrejövő adatsorok kiértékelése már nem képezi minden esetben a tematika szerves részét. Ezért először leendő munkahelyükön kell szembesülniük a létrejött adathalmazok kezelésének és feldolgozásának problémáival. E tananyag publikálásával ezt a hiányt kívántuk pótolni, így az elméleti anyagrész alárendelt szerepet kapott, míg a gyakorlati alkalmazás kiemelt hangsúllyal szerepel. További fontos célunk az adatelőkészítés mellett, hogy a hallgató minél korábban találkozzon több szakterületről származó tudományos problémával, továbbá olyan nagy mennyiségű adathalmazzal, amit már csak számítástechnikai alkalmazás használatával tud kezelni.

A jegyzet a Microsoft Excel program használatán keresztül, valós szakmai példákkal és adatokkal mutat be függvényeket és azok összetett alkalmazásait, melyeket mi, szerzők is rendszeresen használunk munkáink, kutatásaink során. A könyvhöz tartozó elektronikus adatok, illetve fájlok lényeges és szerves részét képezik a tartalomnak. Kiemelten fontosnak tartjuk, hogy a felhasználó az alkalmazandó függvényeket ne a függvényvarázsló használatával alkalmazza, hanem azokat maga írja be, megtanulva a szintaktikát, és egyben felkészülve olyan szoftverek használatára is, melyek magasabb szintű feladatok megoldására is alkalmasak, így igényeknek megfelelően „programozhatók”.

Felmerülhet a kérdés, miért éppen a Microsoft Excelre esett a választás, amikor a szerzők be kívánják mutatni a feladatok megoldásait? A döntés legfőbb oka, hogy a program alapszintű használata középiskolai tananyag, és majdnem minden munkahelyen megtalálható.

Arra a feltételezésre építünk tehát, hogy a középiskolai oktatásból kikerült és a felsőoktatásba felvételt nyert hallgatók valóban rendelkeznek az MS Excel kezeléséhez szükséges alapvető ismeretekkel, így itt csak azon ismeretek kerülnek bemutatásra, amelyek felfrissítése elengedhetetlen a jegyzetben érintett problémák megoldásához. Fontos megemlíteni, hogy a könyvben előforduló függvények bemutatásánál jelentős mértékben

támaszkodtunk az Excel Súgójára, melynek használatáért ezúton is köszönetet mondunk a Microsoft Magyarországnak.

A könyv tíz fejezetből áll, szinte mindegyiket más-más személy írta. Ebből következően a megfogalmazás stílusában vannak kisebb különbségek. Néhány feladathoz nem készült megoldás munkalap, mivel ezekben az esetekben a könyvben szereplő megoldás olyan részletes, hogy külön megoldás munkalap létrehozását nem találtuk szükségesnek. Reményeink szerint az említett tények nem fogják zavarni a felhasználót.

Köszönet illeti a lektorokat, dr. Zibolen Endrét, Zibolen Erzsébetet és Tanos Mártont a kézirat ellenőrzéséért, pontos és gyors munkájukért.

Budapest, 2015. november

a Szerzők

1. Hasznos billentyűkombinációk és cellahivatkozások

A feladat célja:

A Microsoft Excel hasznos billentyűkombinációnak, a cella- és munkalap-hivatkozások módjainak megismertetése.

Új ismeretek:

- Kurzor mozgatása és billentyűkombinációk
- Munkalapok közötti hivatkozások
- Abszolút és relatív cellahivatkozások (távolságmátrix)

Elméleti háttér:

Az egér használata nagyban megkönnyíti a tanulási folyamatot, amikor azonban lépések sorozatát kell végrehajtani, és/vagy az adatok mennyisége, a sorok vagy oszlopok (vagy mind a kettő) száma nagyon nagy, a billentyűzet szinte kizárólagos használata célravezetőbb. A fejezetben olyan billentyűkombinációkat mutatunk be, melyekkel az egérhasználat jelentős része kiváltható.

Az abszolút és relatív cellahivatkozások tárgyköre viszont nem „kényelmi funkció”, hanem olyan anyagrész, melynek ismerete a felhasználó számára elengedhetetlen. Amikor egy cella képletet tartalmaz és ezt a cellatartalmat átmásoljuk egy másik cellába, akkor a tartalomban lévő *relatív* cellahivatkozások meg fognak változni. Amennyiben viszont abszolút cellahivatkozást használunk, \$ „dollárjellel” megjelöljük a sort és/vagy az oszlopot, akkor az oszlop és/vagy a sor azonosítója nem fog változni a másolás során.

Feladatok:

(a) Hasznos billentyűkombinációk:

Billentyűkombináció	Funkció
ALT+F8	A makrók létrehozására, szerkesztésére vagy törlésére szolgáló Makró párbeszédpanel megjelenítése.
ALT+PAGE DOWN	Mozgás a munkalapon egy képernyővel jobbra.
ALT+PAGE UP	Mozgás a munkalapon egy képernyővel balra.
CTRL+A	A teljes munkalap kijelölése.
	Ha a munkalap tartalmaz adatokat, a CTRL+A billentyűparancs kijelöli az aktuális területet. A CTRL+A másodszori lenyomására az alkalmazás kijelöli a teljes munkalapot.
CTRL+B	A félkövér formázás alkalmazása vagy eltávolítása.
CTRL+C	A kijelölt cellák másolása.

CTRL+END	Ugrás a munkalap utolsó cellájára (legalsó felhasznált sor, legtávolabbi felhasznált oszlop). Ha a kurzor a szerkesztőlécen található, ugrás a „szöveg” végére.
CTRL+F	A Keresés és csere párbeszédpanel megjelenítése a Keresés lap kiválasztott állapotában.
CTRL+H	A Keresés és csere párbeszédpanel megjelenítése a Csere lap kiválasztott állapotában.
CTRL+HOME	Ugrás a munkalap elejére
CTRL+I	A dőlt formázás alkalmazása vagy eltávolítása.
CTRL+N	Új, üres munkafüzet létrehozása.
CTRL+NYÍL BILLENTYŰ	Ugrás a munkalapon az aktuális adatterület szélére.
CTRL+O	A fájlok megnyitására vagy megkeresésére szolgáló Megnyitás párbeszédpanel megjelenítése.
CTRL+P	Megjeleníti a Nyomtatás lapot
CTRL+PAGE DOWN	Váltás a munkafüzet következő lapjára.
CTRL+PAGE UP	Váltás a munkafüzet előző lapjára.
CTRL+S	Az aktív fájl mentése az aktuális fájlnevével és fájlformátummal az aktuális helyre.
CTRL+SHIFT+END	A kijelölés bővítése a munkalap legutolsó használatban lévő cellájáig (jobb alsó sarok).
CTRL+SHIFT+NYÍL BILLENTYŰ	A kijelölés bővítése az aktív cella sorának vagy oszlopának utolsó nem üres cellájáig, vagy ha a következő cella üres, a kijelölés bővítése a következő nem üres celláig.
CTRL+SHIFT+P	A Cellák formázása párbeszédpanel megjelenítése a Betűtípus lap kiválasztott állapotában.
CTRL+SZÓKÖZ	A teljes oszlop kijelölése a munkalapon.
CTRL+U	Az aláhúzás alkalmazása vagy eltávolítása
CTRL+V	A vágólap tartalmának beszúrása a kurzor aktuális helyére, és minden kijelölés megszüntetése. Csak akkor elérhető, ha előzőleg egy objektum, szöveg vagy cellatartalom kivágásra vagy másolásra került.
CTRL+X	A kijelölt cellák kivágása.
CTRL+Z	Az utoljára alkalmazott parancs visszavonása vagy az utoljára megadott bejegyzés törlése a Visszavonás parancssal.
F4	Cellahivatkozás rögzítése képlet belsejében
F12	A Mentés másként párbeszédpanel megjelenítése.
HOME	Ugrás a sor elejére a munkalapon.
PAGE DOWN	Mozgás a munkalapon egy képernyővel lefelé.
PAGE UP	Mozgás a munkalapon egy képernyővel felfelé.
SHIFT+F11	Új munkalap beszúrása.
SHIFT+NYÍL BILLENTYŰ	A kijelölés bővítése egy (vagy több) cellával.
SHIFT+SZÓKÖZ	A teljes sor kijelölése a munkalapon.

(b) Munkalapok közötti hivatkozások:

Amennyiben az alapadathalmazunktól eltérő munkalapon kívánunk dolgozni, úgy szükségessé válik annak meghivatkozása.

Ha például a „Munka2” munkalap A1-es cellájában szeretnénk meghivatkozni a „Munka1” munkalap F2-es celláját, azt úgy tehetjük meg, hogy a „Munka2” munkalap A1-es cellájába a =Munka1!F2 képletet írjuk az alábbiak szerint :

A1 → =Munka1!F2

Természetesen egy képletben belül több munkalapot is meghivatkozhatunk. Például, ha a „Munka2” munkalap A1-es cellájában a „Munka1” munkalap F2-es cellájának értékéből kivonjuk a „Munka3” munkalap B4-es cellájának értékét, akkor az alábbi képletet kell a „Munka2” munkalap A1-es cellába írunk:

A1 → =Munka1!F2-Munka3!B4

(c) Abszolút és relatív cellahivatkozások

Egy *cellahivatkozás* adott munkalap celláját vagy tartományát azonosítja. Meghatározza, hogy egy képletben használni kívánt „érték” vagy „értékek” hol található. Az MS Excel az oszlopokra betűkkel, a sorokra sorszámmal hivatkozik, ezek az oszlop- és sorazonosítók. Cellahivatkozás esetén először az oszlopazonosítót, majd a sorazonosítót adjuk meg SZÓKÖZ nélkül. Például az „=B2” hivatkozás a B oszlop és a 2-es sor metszéspontjában található cella értékét adja eredményül.

A *relatív cellahivatkozás* (pl. „=A1”) a képletet tartalmazó és a hivatkozott cella közötti relatív elhelyezkedésen alapul. Ha a képletet tartalmazó cella helye változik, illetve ha a képletet máshová másoljuk, hivatkozása automatikusan igazodik az új helyhez. Alapértelmezés szerint az Excel relatív hivatkozásokat használ. Ha például „=A1” egy relatív hivatkozást a B2 cellából a B3 cellába másoljuk, az eredeti képlet „=A2” képletre módosul.

Az *abszolút cellahivatkozás* (pl. „=\$A\$1”) a képletben mindig – tehát a képlet másolása után is – ugyanazon helyen lévő cellára mutat (esetünkben az A1-re). Az abszolút cellahivatkozás jele: \$. Mivel alapértelmezés szerint az Excel relatív hivatkozást használ, adott esetben a felhasználónak kell beállítani az abszolút cellahivatkozásokat.

Fontos, hogy ha egy tetszőleges hivatkozást tartalmazó cella helye változik – például áthelyezés (nem másolás), oszlop vagy sortörlés következtében – a benne lévő hivatkozás NEM igazodik az új helyhez, azaz változatlan marad.

Relatív és abszolút hivatkozás kombinálható is egy hivatkozáson belül (pl. „=\$A1” vagy „=A\$1”). Amennyiben csak az oszlopazonosító elé kerül a \$ jel, a sorhivatkozás automatikusan igazodik, míg az oszlophivatkozás nem igazodik az új helyhez. (Pl. az „=\$A1” hivatkozást a B2 cellából a C3-ba másolva „=\$A2” az eredmény). Értelemszerűen a sorazonosító rögzítése esetén (pl. „=A\$1”) a sorhivatkozás nem igazodik az új helyhez sem másolás, sem áthelyezés esetén.

Adott területre vagy cellatartományra történő hivatkozás során a bal felső és jobb alsó cellák cellahivatkozásait a „:” jellel kapcsoljuk össze (pl. „=B2:D5”).

Abban az esetben, ha több cellára szeretnénk hivatkozni és a közöttük lévő cellákra nem, elválasztójelként a „;” jelet kell alkalmazni (pl. „=B2;D5”).

Néhány egyszerű példán keresztül szemléltetjük az 1.c pontban bemutatottakat.

$$D4 \rightarrow = \$J\$12 + \$G1 - E\$1 + C4 - \$A\$1$$

D4 cella tartalmát átmásoljuk a G9, L17 és F13 cellákba. A fent bemutatott cellahivatkozások a következőképpen változnak meg.

$$G9 \rightarrow = \$J\$12 + \$G6 - H\$1 + F9 - \$A\$1$$

$$L17 \rightarrow = \$J\$12 + \$G14 - M\$1 + K17 - \$A\$1$$

$$F13 \rightarrow = \$J\$12 + \$G10 - G\$1 + E13 - \$A\$1$$

A következő alfejezetben egy távolságmátrix példáján fogjuk bemutatni az itt leírt ismeretek gyakorlati alkalmazását.

(d) Összefoglaló feladat, távolságmátrix

Példafájl neve: tavmat.xlsx

A feladat célja:

Egy gyakorlati példán keresztül elmélyíteni a rögzített cellahivatkozások használatának ismeretét. Ezt egy távolságmátrix elkészítésével érjük el, amiben ki kell számítani néhány magyarországi város egymástól való távolságát.

Új ismeretek:

- SZUM függvény használata
- Adatok transzponálása

Elméleti háttér (BÁCSATYAI, 2005 alapján):

Az EOVS (Egységes Országos Vetület) 1975-óta használatos vetületi rendszer Magyarországon. Alapfelülete a Nemzetközi Geodéziai Unió által elfogadott UTM67 forgási ellipszoid. Az X tengely pozitív oldala északra mutat, míg az Y tengely pozitív oldala keletre mutat, vagyis a koordinátatengelyek derékszöget zárnak be és tájolása ÉK-i. Az egész ország képe az első koordinátanegyedben található, így az ország minden pontja két pozitív számmal megadható. Az X koordináta értéke minden esetben 400000-nél kisebb, míg az Y koordináta értéke minden esetben nagyobb, mint 400000. A vetületi rendszer egységnyi beosztásának hossza 1 m.

Feladatban felhasznált függvények, funkciók:

- **SZUM függvény:** Egy tartományon belüli értékek összegét adja eredményül.

=SZUM(szám1; [szám2]¹;...)

A **transzponálás** funkció függőleges cellatartományt alakít át vízszintes cellatartománnyá és fordítva. Transzponálás esetén a kijelölt tartomány sorainak száma megegyezik a létrehozott tartomány oszlopainak számával, míg a kijelölt tartomány oszlopainak száma megegyezik a létrehozott tartomány sorainak számával.

Feladatok:

- (a) Egymással összehasonlítva Budapesten vagy az ország más városaiban laknak többen összesen?
- (b) Budapesttől eltekintve a Dunától keletre vagy nyugatra laknak többen?
- (c) Adjuk meg, hogy az egyes városok milyen távolságra fekszenek egymástól, ehhez számítsuk ki a városok távolságmátrixát²!

Megoldás:

(a) feladat:

Nyissuk meg a **tavmat.xlsx** fájlt! Az „ADAT” munkalap tartalmazza az alapadatokat, tehát a városneveket és a vonatkozó EOVS koordinátákat. Az „ADAT” munkalapon az E2 cellába írjuk a SZUM függvényt az alábbi megfelelő cellatartomány hivatkozásával:

E2 → =SZUM(D3:D21)

Tehát Budapest népessége (1735 711 fő) meghaladja a feladatban szereplő többi város össznépességét (1725 867).

(b) feladat:

A G2 cellában számoljuk ki a Dunától nyugatra lévő városok összlakosságát, ehhez a G2 cellába írjuk a következőt:

G2 → =SZUM(D7;D10;D11;D13;D14;D16;D18;D6)

Az eredmény 686 317 fő.

A J2 cellában számoljuk ki a Dunától keletre fekvő városok össznépességét:

¹ A szögletes zárójelben [] lévő argumentumokat nem kötelező megadni.

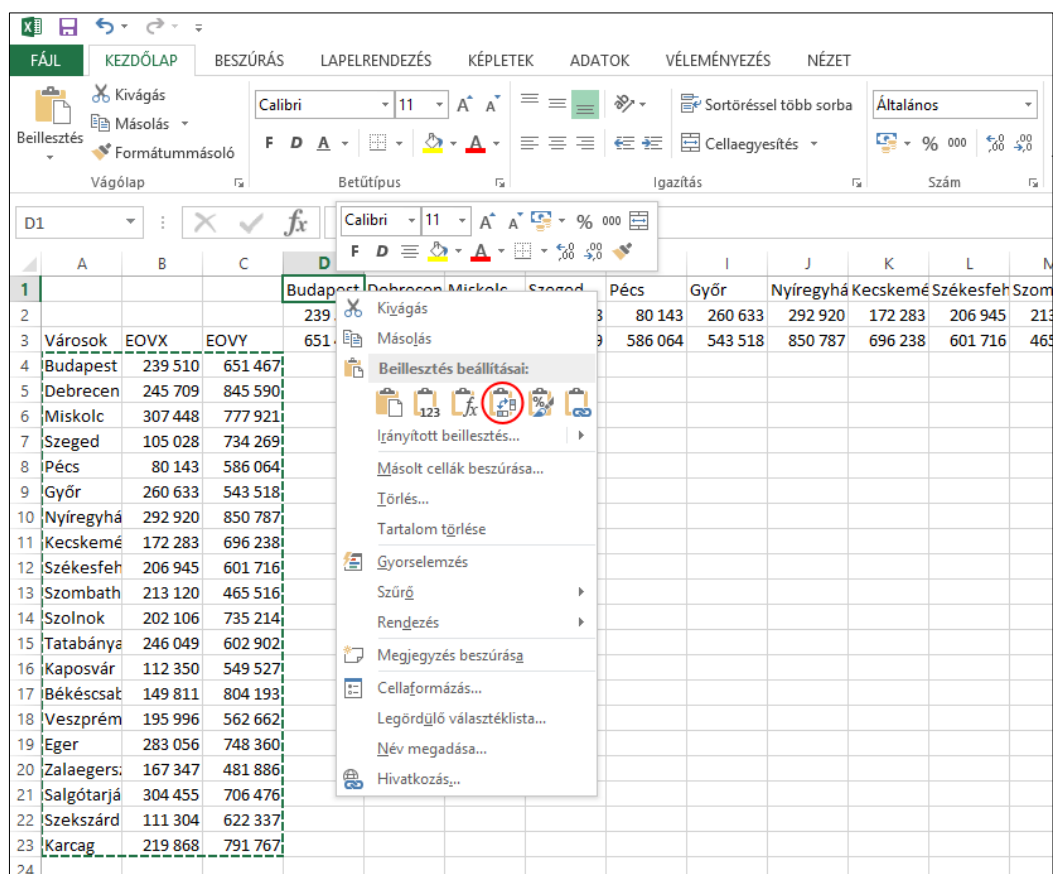
² Olyan mátrix, mely segítségével lehetővé válik, hogy térképi állomány nélkül meghatározzuk két előre definiált térképi pont közötti legrövidebb távolságot.

J2 → = SZUM(D3:D5;D8:D9;D12;D15;D17;D19:D21)

Az eredménye 1 039 550 fő. Tehát keleten laknak többen.

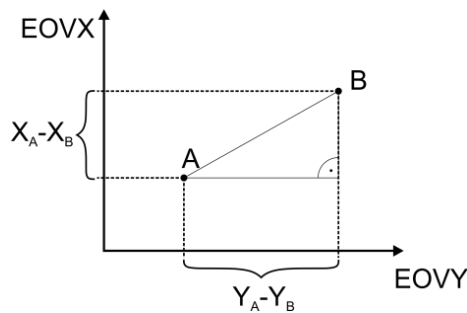
(c) feladat:

A megoldás számára hozzunk létre egy „TAVMAT” nevű munkalapot! Az első, „ADAT” nevű munkalapon található alapadatok segítségével a „TAVMAT” lapon hozzuk majd létre a távolságmátrix kezdősorait és -oszlopait! A „TAVMAT” első három oszlopa tartalmazza sorrendben a városok nevét, majd ezek EOVS, EOVS koordinátáit. Az első három sora pedig ugyanígy a városok nevét, és ezek EOVS, EOVS koordinátáit tartalmazza! Másoljuk át az „ADAT” munkalap A1:C21 tartományát a „TAVMAT” munkalapra, annak A3 celláját választva kezdőpontnak. Ezután a „TAVMAT” lapon jelöljük ki az A4:C23 cellákat, majd CTRL+C billentyű kombinációval másoljuk a vágólapra, végül a D1 cellába kattintsunk jobb egérgombbal és válasszuk a „Beillesztés beállításai” közül a transzponálás funkciót! **(1. ábra)** Egészítsük ki a fejléctet, a C1 cellába írjunk „Városok”, a C2 cellába pedig „EOVS” szöveget.



1. ábra: Adattartomány transzponálása

Számoljuk ki a D4 cellában a D1 és A4 cellákban lévő városok közötti távolságot. Alkalmazzuk a Pitagorasz-tételt a távolság kiszámításához, hiszen az EOVS vetületi rendszer derékszögű koordinátarendszeren alapul **(2. ábra)**.



2. ábra: Két pont közötti távolság meghatározása EOVS koordináta-rendszerben

A D4 cellába írjuk a következőt:

$$D4 \rightarrow = ((\$B4-D\$2)^2+(\$C4-D\$3)^2)^{0,5}$$

A B4 és C4 cellahivatkozás másolása során is csak a B ill. C oszlop mentén mozoghat, ezért a B és a C oszlopot kell rögzíteni \$ jellel, amely tehát így ebben az esetben az oszlopazonosító elé kerül. A D4 képletében ezért írtunk „\$B4”-et és „\$C4”-et. A D2 és D3 cellahivatkozás a D4 képletének másolása során csak sorok mentén mozoghat, ezért itt a sorokat kell rögzíteni, így itt „D\$2”-öt és „D\$3”-at írtunk.

Végül jelöljük ki a D4 cellát, és másoljuk le a tartalmát a W23 celláig! Így létrehoztuk a városok távolságmátrixát.

Összefoglalás: a fejezetben példát adtunk arra, hogy milyen módon lehet az MS Excelben egy adott cella értékére vagy egy cellatartomány értékeire hivatkozni. Bemutattuk a hivatkozások típusait (relatív, abszolút), konkrét cellatartományt transzponáltunk. Végül ezeket felhasználva magyarországi nagyvárosok távolságait számoltuk ki egy távolságmátrixban.

2. Feltételes függvények alkalmazása felszínalatti vízminőség idősorokon

Példafájl neve: B_határérték_2014.xls

Adatforrás: nem nyilvános

A feladat célja:

Felszínalatti víz megfigyelő kutak mintáinak kémiai és fizikai paramétereit tartalmazó adattáblában megkeresni azokat a mérési időpontokat, ahol valamely paraméter értéke eléri vagy meghaladja a „B” szennyezettségi határértéket.

Új ismeretek:

- HA függvény
- DARABTELI függvény
- Feltételes formázás

Elméleti háttér:

„B” szennyezettségi határérték: jogszabályban, vagy ennek hiányában hatósági határozatban megállapított koncentrációérték, melyet a földtani közeg és a felszínalatti víz esetében állapítanak meg (jelenleg: 6/2009. (IV. 14.) KvVM-EüM-FVM együttes rendelet). Azokat a vizeket, illetve azt a földtani közeget tekintjük szennyezettnek, ahol az adott kémiai elem vagy ion eléri a „B” szennyezettségi határértéket. A környezetvédelemben a „B” határérték az egyik leggyakrabban használt érték, amelyhez az adatokat szennyezettségi szempontból hasonlítjuk. Az Európai Unió Víz Keretirányelv (VKI; EC, 2000) célja, hogy az Európai Unió területén minden felszíni és felszínalatti vizet "jó állapotba" hozzon és ezt az állapotot fenntartsa. A VKI céljának elérése érdekében, de a klímaváltozás hatására is egyre inkább felértékelődik a felszíni és a felszínalatti vizek vízminőségének vizsgálata, amely során az értékelést nem közigazgatási határok alapján, hanem vízgyűjtők szerint érdemes végezni. A közigazgatási határok nagyban megnehezítik a vízgyűjtő-szintű vizsgálatot, amint arra MAGYAR ET AL. (2013a, 2013b) is tanulmányai is rávilágítottak.

A feladatban használt adathalmaz egy felszínalatti víz monitoring rendszer vízmintáinak laboratóriumi eredményeit tartalmazza. Ezen adatok alapján vizsgáljuk meg, hogy az adott víztest mennyiben tekinthető szennyezettnek, azaz az adatok mekkora része éri el a jogszabályban meghatározott „B” szennyezettségi határértéket.

Feladatban felhasznált függvények:

- HA függvény:

A függvény egy meghatározott értéket ad eredményül, ha egy megadott feltétel IGAZ értékű, és egy másik értéket, ha a feltétel HAMIS. Például a felszínalatti vízben a nitrát-koncentráció

idősorát vizsgálva a függvény választ tud adni arra, hogy mely időpontokban haladta meg a koncentráció a „B” szennyezettségi határértéket.

=HA(logikai_vizsgalat; [érték_ha_igaz]; [érték_ha_hamis])

logikai_vizsgalat: megadása kötelező. Tetszőleges érték vagy kifejezés, amelynek kiértékelésekor IGAZ vagy HAMIS eredmény adódik.

érték_ha_igaz: megadása nem kötelező. Az az érték, amit arra az esetre írunk elő, amikor a logikai_vizsgalat eredménye IGAZ. Ha például a logikai_vizsgalatban a „B” határértéknél nagyobb nitrát-koncentrációkat keressük és a logikai_vizsgalat eredménye IGAZ, valamint az érték_ha_igaz argumentumba a „szennyezett” karakterláncot adtuk meg, akkor a HA függvény végeredményül a „szennyezett” szöveget adja. Ha a logikai_vizsgalat IGAZ, és az argumentumot üresen hagytuk (tehát a logikai_vizsgalat argumentumot csak egy pontosvessző követi, ami az érték_ha_igaz-nak felel meg), akkor a HA függvény visszatérési értéke 0 (nulla) lesz. Ahhoz, hogy ilyenkor a 0 helyett az IGAZ szó jelenjen meg, használjuk az IGAZ logikai értéket az érték_ha_igaz argumentumban, vagyis a HA(logikai_vizsgalat;IGAZ) változatot.

érték_ha_hamis: megadása nem kötelező. Ha ezt az értéket mégis megadjuk, akkor amennyiben a logikai_vizsgalat argumentum HAMIS eredményt ad, ezt az értéket fogja végeredményként megjeleníteni a HA függvény. Például, ha a logikai_vizsgalatban a „B” határértéknél nagyobb nitrát-koncentrációkat keressük és a logikai_vizsgalat eredménye HAMIS, és az érték_ha_hamis argumentumba a „nem szennyezett” karakterláncot adtuk meg, akkor a HA függvény a „nem szennyezett” szöveget adja eredményül. Ha a logikai_vizsgalat eredménye HAMIS, és az érték_ha_hamis argumentumot nem adtuk meg, akkor a HA függvény a HAMIS logikai értéket adja vissza. Ha a logikai_vizsgalat HAMIS, és a HA függvényben az érték_ha_igaz argumentum után nincs pontosvessző, a HA függvény visszatérési értéke 0 (nulla) lesz.

- **DARABTELI függvény:**

A DARABTELI függvény egy cellatartomány egy megadott feltételnek megfelelő celláit számolja meg. Használhatjuk például arra, hogy egy adott értéknél kisebb vagy nagyobb értékeket tartalmazó cellákat számoljunk meg. (Több feltétel szerinti keresés esetén használható a DARABHATÖBB függvény.

=DARABTELI (tartomány; feltétel(ek))

tartomány: megadása kötelező, mely egy vagy több összeszámolandó cellából áll; tartalmazhat számokat, neveket, tömböket vagy számokat tartalmazó hivatkozásokat. A függvény figyelmen kívül hagyja a tartomány szövegértéket tartalmazó, valamint az üres celláit.

feltételek: legalább egy feltétel megadása kötelező. A feltétel lehet az összeszámolandó cellákat meghatározó számként, kifejezésként, cellahivatkozásként vagy szöveggként megadott feltétel. Ez például a következő formákban adható meg: 32, "32", ">32", "nitrát" vagy B4.

- A feltételek argumentumban használhatunk helyettesítő karaktereket, tehát a kérdőjelet (?) és a csillagot (*). A kérdőjel egy darab tetszőleges karaktert, a csillag pedig bármilyen hosszú tetszőleges karaktersort helyettesíthet. Amennyiben pedig

ténylegesen kérdőjelet vagy csillagot szeretnénk adatként keresni, akkor egy tilde karaktert (~) kell írunk a keresett karakter elé.

- A függvény nem tesz különbséget kis- és nagybetű között.

Feladatok:

- (a) Határozzuk meg, hogy a vizsgált időszakban a monitoring kutakban mért alábbi paraméterek közül hány esetben és mely paraméterek érték el, vagy haladták meg a környezetvédelmi „B” határértéket!

- kadmium, higany, nikkel, cink, ólom, szulfát, nitrát

- (b) Keressük meg azokat a mintavételi időpontokat, ahol az előző feladatban szereplő paraméterek a „B” szennyezettségi határérték 80%-ánál nagyobbak, de nem haladják meg a határértéket!

- Azoknál az időpontoknál, ahol a feltétel teljesül, a függvény visszatérési értéke legyen „magas koncentráció, de nem szennyezett” szöveg. Ha a mért érték eléri vagy meghaladja a „B” határértéket, akkor a függvény azt írja ki, hogy „szennyezett”, illetve ha nem éri el a 80%-ot, akkor pedig azt, hogy „nem szennyezett”.

- (c) Feltételes formázás használatával emeljük ki azokat a cellákat, ahol a „szennyezett”, illetve a „közepesen szennyezett” visszatérési értéket kaptuk az előző feladatban!

Megoldás:

(a) feladat:

Nyissuk meg a **B_határérték_2014.xls** fájlt! Az „ADAT” munkalap tartalmazza mintavételi eseményenként az egyes fizikai és kémiai paraméterek vízmintákban mért értékeit, valamint a kút nevét és a mintavétel időpontját. A „főinok” mg/l-ben, a nyomelemek ppb-ben vannak feltüntetve.

A „B_szenyezettségi határérték” munkalap tartalmazza az egyes paraméterekre vonatkozó határértékeket. Figyeljünk arra, hogy az adatok különböző mértékegységekben lettek megadva.

Az feladat megoldásához a DARABTELI függvényt fogjuk alkalmazni. Elsőként a szulfát példáján nézzük meg, hogy hány esetben érte el vagy haladta meg a szennyezettségi határértéket.

A számításokat mindig az adott oszlop utolsó sora alatt végezzük el, a szulfát esetében a V2266-os cellában. A függvény tartományának adjuk meg az összes szulfátra vonatkozó adatot, ebben az esetben a V2:V2264 cellatartományt. A függvény ezeket a cellákat fogja megszámlálni a feltételek szerint. A szulfát „B” szennyezettségi határértéke 250 mg/l (lásd: „B_szenyezettségi határérték” munkalap). Ez az érték lesz a függvény feltétele: „<=250”. A feladatra alkalmazott függvény:

V2266 → = DARABTELI(V2:V2264;">=250")

Az „ADAT” munkalapon lépünk a V2266 cellába és írjuk be ezt a függvényt. Eredményül 11-et kapunk, amit azt jelenti, hogy 11 esetben érte el vagy haladta meg a szulfáttartalom a szennyezettségi határértéket.

Másoljuk ezt a képletet egyesével a feladatban vizsgálandó paraméterek adott oszlopaiba, a 2266. sorba és írjuk át a feltételben szereplő értéket az adott paraméter szennyezettségi határértékére. Eredményül megkapjuk az összes vizsgálni kívánt valószínűségi változó esetében a határértéket elérő, vagy az azt meghaladó mérések számát.

(b) feladat:

A megoldáshoz először készítsük el a fejléct. Írjuk be az „ADAT” munkalapon az AZ1 cellától kezdve rendre a BF1 celláig a Kadmium_b, Higany_b, Nikkel_b, Cink_b, Ólom_b, Szulfát_b, Nitrát_b szavakat az első sorba, mivel ezekbe az oszlopokba fogjuk kiíratni a feladatban kért eredményeket.

A feladat megoldásához HA függvényt alkalmazunk. Először gyakorlásként nézzünk egy egyszeres (egy logikai vizsgálatot tartalmazó) HA függvényt, majd ennek megértése után egy bonyolultabb, kétszeres HA függvény segítségével oldjuk meg a feladatot.

A függvény első argumentumában adjuk meg a feltételt. Ha arra vagyunk kíváncsiak, hogy mikor volt a kadmium esetén a mért érték a határérték (1 mg/l) 80%-a alatt, akkor legyen a feltétel:

„AO2<1*0,8”

Az AO2 cella tartalmazza a mért kadmium-koncentrációt. 1 mg/l a kadmium szennyezettségi határértéke, és ennek 0,8 szorosa a határérték 80%-a. Tehát amennyiben a mért érték 0,8 alatt van, akkor a függvény eredménye IGAZ lesz, amikor azonban e felett, akkor HAMIS. Ezek alapján a határérték 80%-át így számoljuk ki:

„AO2*0,8”

Mivel el kell különítenünk a 80% és 100% közötti, illetve a 100% fölötti eredményeket, ezért egy összetett HA függvényt kell alkalmazunk. A feladatnak megfelelően az *érték_ha_igaz* argumentum legyen: „nem szennyezett”,

=HA(AO2<1*0,8;"nem szennyezett";[érték_ha_hamis])

az *érték_ha_hamis* argumentumba pedig beillesztünk egy második HA függvényt, mely azt vizsgálja, mikor nagyobb a mért kadmium érték a szennyezettségi határnál (tehát a 100%-nál), és mikor van az alatt. Tehát a második HA függvény feltételeként megadjuk, hogy mikor nagyobb a mért érték a kadmium határértékénél (1 mg/l).

„AO2>=1”

A második HA függvény *érték_ha_igaz* kimenete, amikor 1-nél nagyobb értékről van szó, tehát a határérték 100%-ánál nagyobbbról, legyen „szennyezett”, HAMIS esetben pedig (amikor az érték 100% alatt van) legyen „közepesen szennyezett”. Ezek alapján a második HA függvény a következő:

$HA(AO2 \geq 1; "szennyezett"; "közepesen szennyezett")$

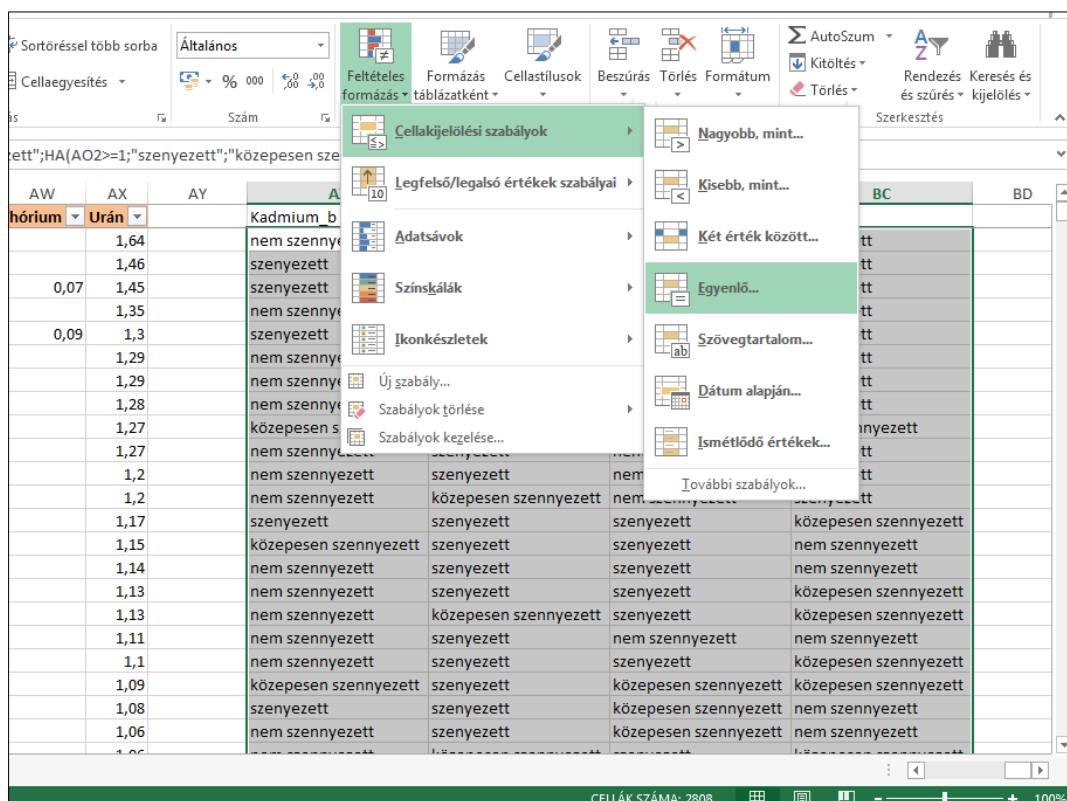
Ezt írjuk az első HA függvény *érték_ha_hamis* argumentumába, így a teljes függvény a következőképpen alakul:

$AZ2 \rightarrow = HA(AO2 < 1 * 0,8; "nem szennyezett"; HA(AO2 \geq 1; "szennyezett"; "közepesen szennyezett"))$

Lépünk az AZ2 cellába és illesszük be a fenti függvényt. Másoljuk át a függvényt a feladatban létrehozott új oszlopok második soraiba az AZ2-től a BF2 celláig. Javítsuk ki a cellahivatkozásokat, hogy mindegyik függvény az adott paraméterre vonatkozó cellákra hivatkozzon. Másoljuk le a függvényeket a munkalap 2262. soráig.

(c) feladat:

Jelöljük ki az AZ1:BD2262 tartományt, majd menjünk a „Kezdőlap” szalagra és kattintsuk a „Feltételes formázás” opcióra, majd a „Cellakijelölés szabályok” menüre, végül ezen belül az „Egyenlő” opcióra (3. ábra).



3. ábra: Feltételes formázás

Baloldalon írjuk be a keresendő értékhez a „szennyezett” szót, jobboldalon hagyjuk az alapértelmezett „Piros kitöltő szín, sötétvörös szöveggel” opciót, majd kattintsunk az OK

gombra. Az Excel megjelöli azokat a cellákat, ahol a HA függvény eredménye „szennyezett” volt. Ismételjük meg a műveletsort úgy, hogy a „közepesen szennyezett” cellák „Sárga kitöltő színnel és sötét sárga szöveggel” jelenjenek meg (4. ábra).

AY	AZ	BA	BB	BC	BD
	Kadmium_b	Higany_b	Nikkel_b	Nitrát_b	
	nem szennyezett	szennyezett	közepesen szennyezett	szennyezett	
	szennyezett	szennyezett	nem szennyezett	szennyezett	
	szennyezett	szennyezett	szennyezett	szennyezett	
	nem szennyezett	szennyezett	szennyezett	szennyezett	
	szennyezett	nem szennyezett	nem szennyezett	szennyezett	
	nem szennyezett	közepesen szennyezett	közepesen szennyezett	szennyezett	
	nem szennyezett	szennyezett	közepesen szennyezett	szennyezett	
	nem szennyezett	közepesen szennyezett	nem szennyezett	szennyezett	
	közepesen szennyezett	szennyezett	szennyezett	nem szennyezett	
	nem szennyezett	szennyezett	nem szennyezett	szennyezett	
	nem szennyezett	szennyezett	nem szennyezett	szennyezett	

4. ábra: A feltételes formázás eredménye

Összefoglalva: a feladat bemutatta, hogy egy felszínalatti vízminta-idősorból a DARABTELI és a HA függvények segítségével néhány egyszerű lépéssel milyen hasznos információk nyerhetők ki. A gyakorlat során a több mint 2200 mintavételi időpontot tartalmazó adathalmazból gyorsan megtudhattuk, hogy az egyes paraméterek esetében hányszor van adathiány, illetve a fejezet ismertetett egy egyszerű módszert, amely segítségével meghatározható, hogy az egyes paraméterek hányszor közelítik meg, érik el, illetve haladják meg a szennyezettségi határértéket.

3. Szöveg- és dátumkezelő függvények

Példafájl neve: FVK.xlsx

Adatforrás: nem nyilvános

A feladat célja:

Bemutatni a szöveg- és dátumkezelő függvények használatát egy gyakori feladat példáján, amelyben szöveggént tárolt dátumot kell dátumformátumúvá alakítani.

Új ismeretek:

- Szövegkezelő függvények használata
- Dátumkezelő függvények használata

Elméleti háttér:

A környezeti rendszerek fizikai, kémiai és biológiai jellemzőinek vizsgálata kiemelten fontos a környezettudományokban. Egy ilyen rendszer tulajdonságainak részletes megismeréséhez egy jól tervezett monitoringrendszer üzemeltetése, és ennek részeként többszöri mintavétel szükséges (MAGYAR ET AL., 2013a). A mintavételek gyakoriságának megválasztása jelentősen függ a vizsgálni kívánt folyamat időbeli (KOVÁCS ET AL. 2012a) és/vagy térbeli változékonyságától (KOVÁCS ET AL., 2012b). Előfordulhat, hogy nagyon rövid időközönként kell mintát venni a rendszerből, ilyenkor általában a mintavételt egy arra alkalmas készülék végzi, amely a mért értékeket folyamatosan regisztrálja, tárolja (jelen példában a felszínalatti víz fajlagos vezetőképességét). Az adatokat azonban gyakran olyan formátumban rögzíti a készülék, amely közvetlenül nem alkalmas a MS Excelben (vagy más programokban) történő elemzéshez. Előfordulhat például, hogy a mintavétel idejét vagy a mért értékeket szöveg formátumban tudjuk csak beimportálni egy Excel munkalapra, ebben a formátumban azonban közvetlenül nem lehet ezeket bevonni a szükséges számításokba. Többek között ilyen jellegű problémák megoldását segítik az MS Excel szöveg- és dátumkezelő függvényei.

A MS Excel dátumkezelése:

Az MS Excel a dátumokat számként tárolja. Az 1-es számhoz az 1900.01.01 0:00 időpontot rendeli a program. 1 nap növekedés (csökkenés) 1, 1 óra 1/24, 1 perc 1/1440, 1 másodperc pedig 1/86400 egységnyi növekedést (csökkenést) jelent a számértékben. Vegyünk egy példát: ha a 30,5-öt dátum formátumúvá alakítjuk, akkor 1900. január 30. 12:00-t kapunk eredményül.

Feladatban felhasznált függvények:

- **JOBB, BAL és KÖZÉP függvény:**

Egy szöveg megadott számú karaktereit adja eredményül jobbról, balról, illetve egy megadott kezdő sorszámtól kezdve.

=**JOB**B(szöveg; [hány_karakter])

=**BAL**(szöveg; [hány_karakter])

=**KÖZÉP**(szöveg; honnantól; [hány_karakter])

- **HOSSZ függvény:** egy szöveg karaktereinek számát adja eredményül.

=**HOSSZ**(szöveg)

- **ÉRTÉK függvény:** egy számot megjelenítő szöveget számmá alakít át.

=**ÉRTÉK**(szöveg)

szöveg: megadása kötelező. Az a karakterlánc, amelynek bizonyos karaktereit ki szeretnénk nyerni, át szeretnénk alakítani.

hány_karakter: megadása nem kötelező. A kinyerni kívánt karakterek száma. Amennyiben ezt az argumentumot üresen hagyjuk, feltételezett értéke 1 lesz, továbbá ha a megadott érték meghaladja a szöveg hosszát, akkor a teljes szöveget adja eredményül a függvény.

honnantól: megadása kötelező. Az első kinyerni kívánt karakter sorszáma.

- **DÁTUMÉRTÉK függvény:**

Egy szöveggént megadott dátumot olyan számmá alakít át, amely MS Excel dátum- és időértékben adja meg a dátumot.

=**DÁTUMÉRTÉK**(dátum_szöveg)

dátum_szöveg: dátumot jelölő szöveg formátumú karakterlánc.

- **DÁTUM függvény:**

A megadott év, hónap és napnak megfelelő, MS Excel dátum- és időérték formátumú dátumot adja eredményül.

=**DÁTUM**(év;hónap;nap)

év: megadása kötelező. A dátumban az évet jelölő szám.

hónap: megadása kötelező. A dátumban a hónapot jelölő szám.

nap: megadása kötelező. A dátumban a napot jelölő szám.

- **ÉS függvény:** IGAZ értéket ad, ha minden argumentuma IGAZ; HAMIS értéket ad, ha legalább egy argumentuma HAMIS.

=**ÉS**(logikai₁;[logikai₂];...)

logikai₁: megadása kötelező. Az első vizsgálandó feltétel.

logikai₂: megadása nem kötelező. A második vizsgálandó feltétel.

Egyéb, gyakran használt szöveg- és dátumkezelő függvények leírása:

- **ÖSSZEFŰZ függvény:** Több szövegdarabot egyetlen szöveggé alakít.

=ÖSSZEFÜZ(szöveg₁; szöveg₂;...)

- **ÉV függvény:** a dátumértéknek megfelelő évet adja eredményül.
=ÉV(dátumérték)
- **HÓNAP függvény:** a dátumértéknek megfelelő hónapot adja eredményül (1-12).
=HÓNAP(dátumérték)
- **NAP függvény:** a dátumértéknek megfelelő napot adja eredményül (1-31).
=NAP(dátumérték)
- **ÓRA függvény:** az időértéknek megfelelő órát adja eredményül (0-23).
=ÓRA(időérték)
- **PERCEK függvény:** az időértéknek megfelelő percet adja eredményül (0-59).
=PERCEK(időérték)
- **MPERC függvény:** az időértéknek megfelelő másodpercet adja eredményül (0-59).
=MPERC(időérték)
- **MA függvény:** az aktuális dátumot adja eredményül dátum formátumban.
=MA()
- **MOST függvény:** az aktuális dátumot és időpontot adja eredményül dátum és idő formátumban.
=MOST()

Feladatok:

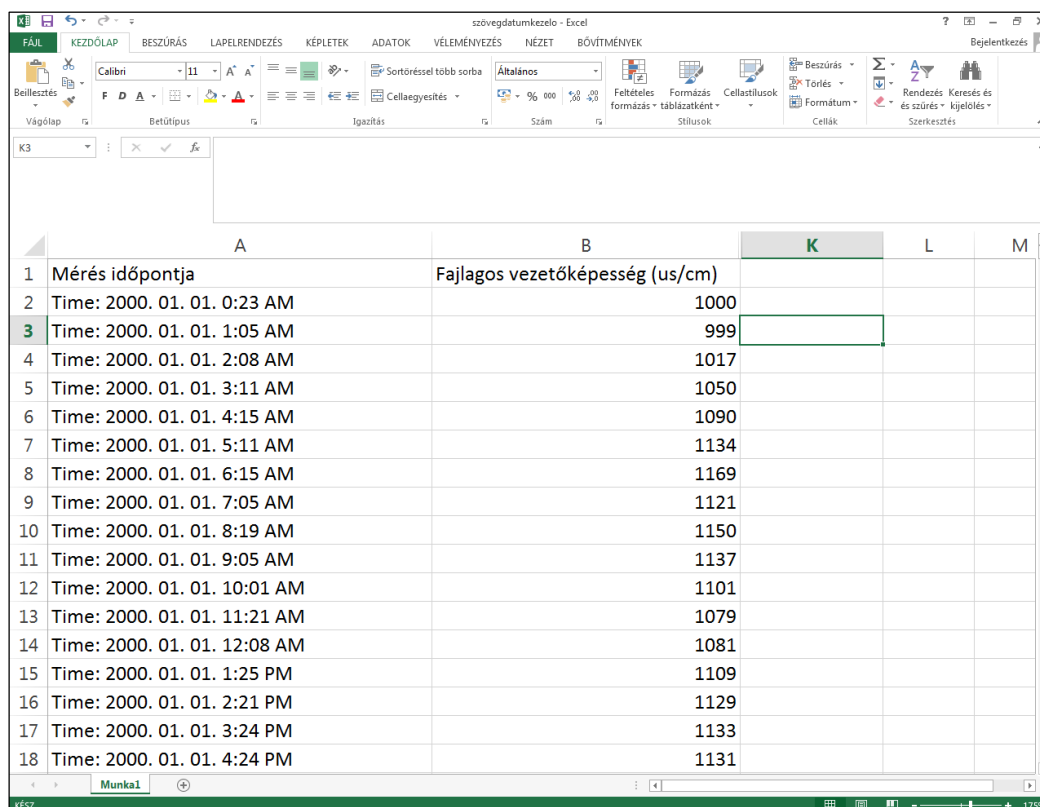
- Nyissuk meg az FVK.xlsx példafájlt! Értelmezzük a mérőműszer által rögzített adatokat! Vizsgáljuk meg, hogy az egyes cellák milyen formátumúak!
- Az A oszlopban tárolt mérési időpontokból írassuk ki szám formátumban a C, D, illetve E oszlopokban a mérések évét, hónapját, napját, majd ezt követően az F oszlopban hozzuk létre ezen adatokból a mérések dátumait! A C, D és E oszlopokat lássuk el az alábbi fejléccel: „Év”, „Hónap”, „Nap”!
- A G és H oszlopokban írassuk ki a mérések idejének óráit 12 (0-11), illetve 24 órás (0-23) formátumban, az I oszlopban pedig a mérések időpontjának percei szerepeljenek! A G, H és I oszlopokat lássuk el az alábbi fejlécekkel: „Óra (12)”, „Óra (24)”, „Perc”!
- A J oszlopban hozzuk létre a mérések pontos időpontjait dátum formátumban (óra, perccel is) a korábban kiszámított értékek felhasználásával! Az oszlopot lássuk el az alábbi fejléccel: „Mérés pontos időpontja”!

- (e) A K oszlopban hozzuk létre a mérések pontos időpontjait dátum formátumban (óra, perccel is) úgy, hogy csak az A oszlopra hivatkozunk a megoldást adó függvényben! Az oszlopot lássuk el az alábbi fejléccel: „Mérés pontos időpontja egy függvénnyel”!

Megoldás:

(a) feladat:

A példafájl 2 oszlopot tartalmaz. Az első a mérések időpontjait szöveg formátumban (mivel szöveg típusú karaktert is tartalmaz a cella), a második pedig a regisztrált értékeket, jelen esetben a fajlagos vezetőképesség-értékeket szám formátumban tartalmazza. Megfigyelhetjük, hogy (alapértelmezésben) a szöveg tartalmú cellákat balra, míg a számformátumú cellákat jobbra zárja a MS Excel (5. ábra).



	A	B	K	L	M
1	Mérés időpontja	Fajlagos vezetőképesség (us/cm)			
2	Time: 2000. 01. 01. 0:23 AM	1000			
3	Time: 2000. 01. 01. 1:05 AM	999			
4	Time: 2000. 01. 01. 2:08 AM	1017			
5	Time: 2000. 01. 01. 3:11 AM	1050			
6	Time: 2000. 01. 01. 4:15 AM	1090			
7	Time: 2000. 01. 01. 5:11 AM	1134			
8	Time: 2000. 01. 01. 6:15 AM	1169			
9	Time: 2000. 01. 01. 7:05 AM	1121			
10	Time: 2000. 01. 01. 8:19 AM	1150			
11	Time: 2000. 01. 01. 9:05 AM	1137			
12	Time: 2000. 01. 01. 10:01 AM	1101			
13	Time: 2000. 01. 01. 11:21 AM	1079			
14	Time: 2000. 01. 01. 12:08 AM	1081			
15	Time: 2000. 01. 01. 1:25 PM	1109			
16	Time: 2000. 01. 01. 2:21 PM	1129			
17	Time: 2000. 01. 01. 3:24 PM	1133			
18	Time: 2000. 01. 01. 4:24 PM	1131			

5. ábra: A szöveg tartalmú cellákat balra, a számformátumú cellákat jobbra zárja a MS Excel

(b) feladat:

A feladat megoldásához először a szükséges információt tartalmazó karakterláncokat kell kinyernünk, majd ezeket szám formátumúvá alakítanunk. Vizsgáljuk meg, hogy a szükséges értékek hol találhatóak az A oszlopban tárolt szövegekben! Az éveket a 7-10., a hónapokat 13-14., míg a napokat 17-18. karakterek jelölik. Mivel ezen karakterek nem a karakterlánc szélén helyezkednek el, így a KÖZÉP függvényt kell használnunk mind a három esetben. A függvény eredménye azonban szöveg formátumú, amelyet az ÉRTÉK függvénnyel alakíthatunk számmá. Ezek alapján az alábbi függvényeket írjuk a C2, D2 és E2 cellákba:

C2 → =ÉRTÉK(KÖZÉP(A2;7;4))

D2 → =ÉRTÉK(KÖZÉP(A2;13;2))

E2 → =ÉRTÉK(KÖZÉP(A2;17;2))

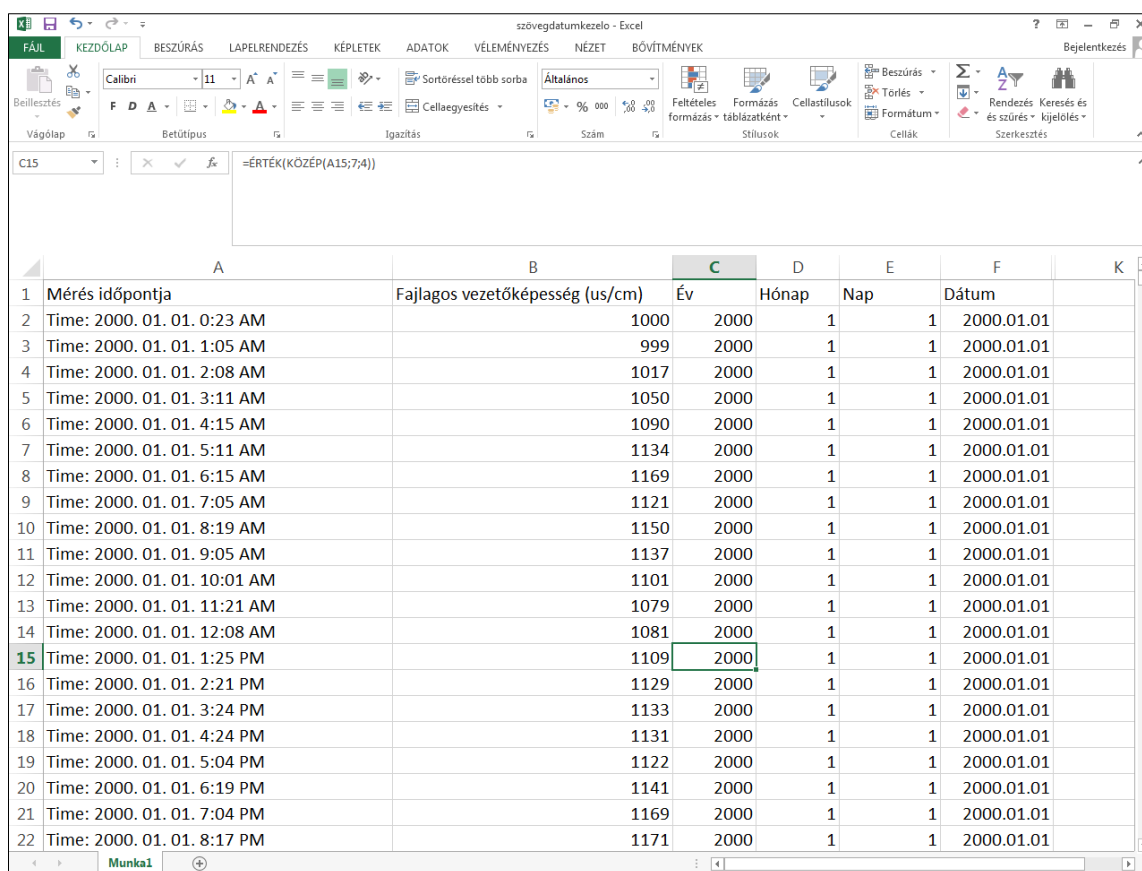
A következő részfeladatunk, hogy az F oszlopban hozzuk létre a mérések dátumait. Ehhez használhatjuk a DÁTUM függvényt, úgy hogy a függvény argumentumában az évet, a hónapot illetve a napot tartalmazó cellákra hivatkozunk:

F2 → =DÁTUM(C2;D2;E2)

Egy másik lehetséges megoldást nyújt a DÁTUMÉRTÉK függvény, mely egy szöveggént tárolt időpont-megjelölést dátum formátumúvá alakít. Ebben az esetben az egész dátumot (a 7-18. karakterig) kinyerjük a mérés időpontját tartalmazó karakterláncból a KÖZÉP függvénnyel, és ezt a DÁTUMÉRTÉK függvénnyel dátumértékké alakítjuk, majd a Kezdőlap menüpontban a cellát dátum formátumúvá formázzuk:

F2 → =DÁTUMÉRTÉK(KÖZÉP(A2;7;13))

Ezt követően másoljuk a C2:F2 cellák függvényeit az adatsor utolsó soráig, majd hozzuk létre a kért fejléct (6. ábra).



	A	B	C	D	E	F	K
	Mérés időpontja	Fajlagos vezetőképesség (us/cm)	Év	Hónap	Nap	Dátum	
1							
2	Time: 2000. 01. 01. 0:23 AM	1000	2000	1	1	2000.01.01	
3	Time: 2000. 01. 01. 1:05 AM	999	2000	1	1	2000.01.01	
4	Time: 2000. 01. 01. 2:08 AM	1017	2000	1	1	2000.01.01	
5	Time: 2000. 01. 01. 3:11 AM	1050	2000	1	1	2000.01.01	
6	Time: 2000. 01. 01. 4:15 AM	1090	2000	1	1	2000.01.01	
7	Time: 2000. 01. 01. 5:11 AM	1134	2000	1	1	2000.01.01	
8	Time: 2000. 01. 01. 6:15 AM	1169	2000	1	1	2000.01.01	
9	Time: 2000. 01. 01. 7:05 AM	1121	2000	1	1	2000.01.01	
10	Time: 2000. 01. 01. 8:19 AM	1150	2000	1	1	2000.01.01	
11	Time: 2000. 01. 01. 9:05 AM	1137	2000	1	1	2000.01.01	
12	Time: 2000. 01. 01. 10:01 AM	1101	2000	1	1	2000.01.01	
13	Time: 2000. 01. 01. 11:21 AM	1079	2000	1	1	2000.01.01	
14	Time: 2000. 01. 01. 12:08 AM	1081	2000	1	1	2000.01.01	
15	Time: 2000. 01. 01. 1:25 PM	1109	2000	1	1	2000.01.01	
16	Time: 2000. 01. 01. 2:21 PM	1129	2000	1	1	2000.01.01	
17	Time: 2000. 01. 01. 3:24 PM	1133	2000	1	1	2000.01.01	
18	Time: 2000. 01. 01. 4:24 PM	1131	2000	1	1	2000.01.01	
19	Time: 2000. 01. 01. 5:04 PM	1122	2000	1	1	2000.01.01	
20	Time: 2000. 01. 01. 6:19 PM	1141	2000	1	1	2000.01.01	
21	Time: 2000. 01. 01. 7:04 PM	1169	2000	1	1	2000.01.01	
22	Time: 2000. 01. 01. 8:17 PM	1171	2000	1	1	2000.01.01	

6. ábra: A (b) feladat végeredménye

(c) feladat:

Az előző feladathoz hasonlóan a mérési időpontok órájának meghatározásához a KÖZÉP függvényt kell használnunk. Fontos észrevennünk azonban, hogy az A oszlopban az egyjegyű óraszámok előtt nem szerepel 0, így ezek esetében 1, míg a kétjegyűeknél 2 karakter tartalmazza az órát. Ez egyben azt is jelenti, hogy a kétjegyű óraszámot tartalmazó cellák karakterlánc 1 karakterrel hosszabb (28), mint az egyjegyűeket tartalmazók (27). Abban az esetben, amikor a karakterlánc hossza 27, akkor 1, egyébként pedig 2 karaktert kell kivennünk a KÖZÉP függvénnyel és értékké alakítanunk. A logikai vizsgálathoz a 3. fejezetben bemutatott HA függvény használható:

G2 → =HA(HOSSZ(A2)=27;ÉRTÉK(KÖZÉP(A2;21;1));ÉRTÉK(KÖZÉP(A2;21;2)))

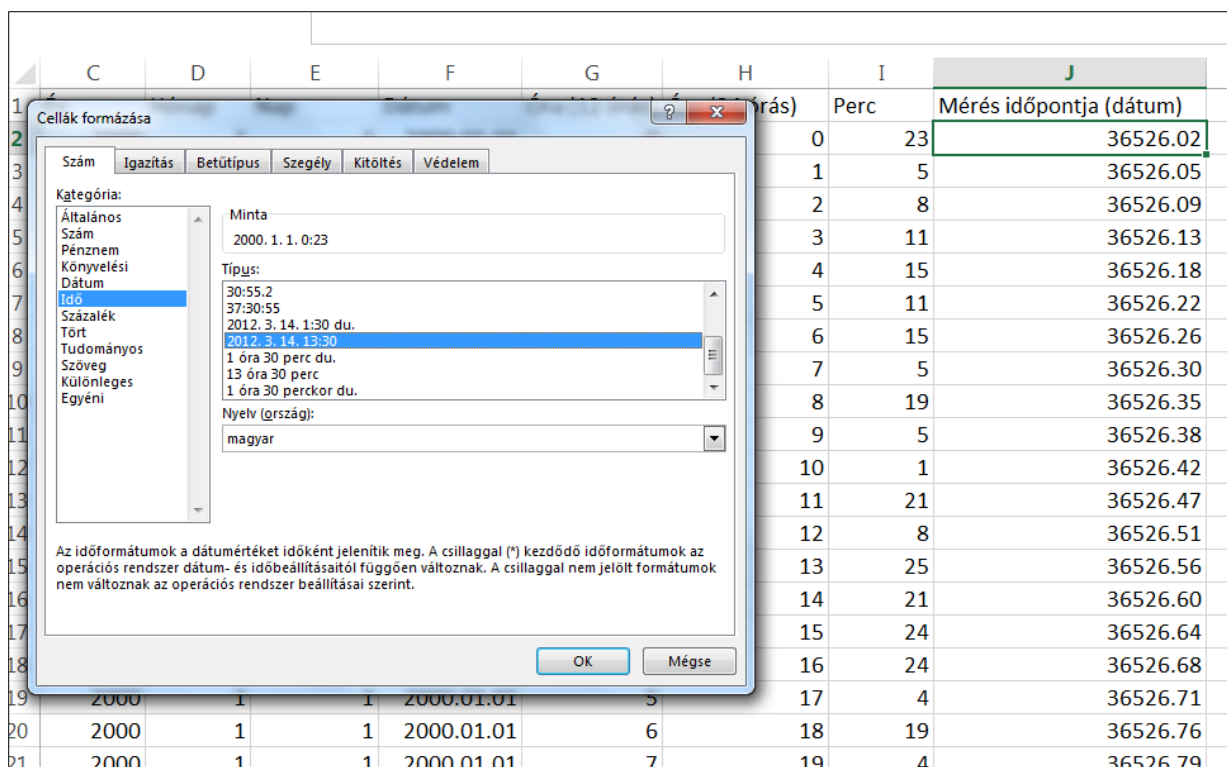
A H oszlopban 24 órás formában kell megadnunk az órákat, vagyis amennyiben a mérés időpontjában „PM” (délután) szerepel, akkor 12 órát hozzá kell adnunk a korábban kiszámolt óra-értékhez. A „PM” szót, mivel az eredeti karakterlánc végén található, legegyszerűbben a JOBB függvénnyel tudjuk kinyerni, jobbról 2 karaktert kell kivennünk a karakterláncból. A feladat megoldása a következő:

H2 → =HA(JOBB(A2;2)=\"PM\";G2+12;G2)

Az órák mellett a percekre is szükségünk van a mérés pontos idejének létrehozásához. Egyjegyű óraszámok esetében a 23-24., míg kétjegyűeknél a 24-25. karakterek tartalmazzák a perceket. A feladat megoldása így az órák kiszámításának módjával megegyezik, csupán a KÖZÉP függvények argumentumait kell módosítanunk:

I2 → =HA(HOSSZ(A2)=27;ÉRTÉK(KÖZÉP(A2;23;2));ÉRTÉK(KÖZÉP(A2;24;2)))

Ezt követően másoljuk a G2:I2 cellák tartalmát az adatsor utolsó soráig, majd hozzuk létre a kért fejléctet (7. *ábra*).



8. ábra: Dátumformátum módosítása

(e) feladat:

Természetesen a mérési időpontok dátummá alakításához nem szükséges, hogy a részadatokat, amint eddig tettük, külön oszlopokban számoljuk ki, hiszen a feladat egy lépésben is megoldható a korábbi függvények és az ÉS függvény együttes használatával:

K2 →

=DÁTUM(ÉRTÉK(KÖZÉP(A2;7;4));ÉRTÉK(KÖZÉP(A2;13;2));ÉRTÉK(KÖZÉP(A2;17;2)))+HA(ÉS(JOBB(A2;2)=""PM";HOSSZ(A2)=27);ÉRTÉK(KÖZÉP(A2;21;1))/24+0.5;HA(ÉS(JOBB(A2;2)=""PM";HOSSZ(A2)=28);ÉRTÉK(KÖZÉP(A2;21;2))/24+0.5;HA(ÉS(JOBB(A2;2)=""AM";HOSSZ(A2)=27);ÉRTÉK(KÖZÉP(A2;21;1))/24;ÉRTÉK(KÖZÉP(A2;21;2))/24)))+(HA(HOSSZ(A2)=27;ÉRTÉK(KÖZÉP(A2;23;2));ÉRTÉK(KÖZÉP(A2;24;2)))/1440)

Végezetül másoljuk a K2 cella tartalmát az adatsor utolsó soráig, majd hozzuk létre a kért fejléceket!

Összefoglalva: a feladat egy példán keresztül bemutatta, hogyan lehet szöveg- és dátumkezelő függvények használatával egy mérőműszer által rögzített, ám szöveggént tárolt időpontokat dátum formátumúvá alakítani.

4. Adatrendeztés és importálás, egy informatikai példa

Példafájl neve: ACCT.txt

A feladat célja:

Egy adatforgalmat mérő műszer által létrehozott „rendezetlen” szöveges információhalmaz feldolgozhatatlan adatelemzési szempontból. A feladat, illetve annak megoldása azt kívánja bemutatni, hogy egy ilyen típusú fájl importálása után milyen problémákkal szembesülhetünk a feldolgozás folyamata során, ami egyúttal kiterjed a fájl importálásától a hasznos cellatartalom kinyeréséig. Összefoglalóan azt mondhatjuk: amennyiben a kapott fájl információtartalmából egy ábrát tudunk készíteni, a feladatot megoldottuk, mert adataink ekkor számként (és nem szöveg formátumban) jelennek meg.

Új ismeretek:

- Szöveges tartalom kezelése függvényekben
- Tartalom összefűzése [&]
- Tartomány másolása

Elméleti háttér:

Mérőeszközökben levő programok sok esetben szöveges (*.txt) formátumban (és akár „rendezetlenül”, egy meghatározott mintázatban) adják meg kimeneti adataikat. Ezeket kezelni kell, ki kell belőlük nyerni a szükséges és feldolgozható információt. A példafájl egy számítógépes labor 100 órás ki- és bemenő adatforgalmát mutatja be. A kapott információhalmaz szerkezetét a megoldásban mutatjuk be.

Feladatban felhasznált függvények (megtalálhatók a 2. és 3. fejezetben):

- **ÖSSZEFŰZ**
- **HOSSZ**
- **ÉV**
- **HÓNAP**
- **NAP**
- **ÓRA**
- **PERC**
- **MPERC**
- **DÁTUMÉRTÉK**
- **HA**

Feladatok:

- (a) Adatimportálás.
- (b) Dátumok és mérési időpontok kinyerése, azok összefűzése és formázása.
- (c) Adathalmazban levő esetleges elcsúszások javítása.
- (d) Az adatok mértékegységének figyelembe vétele és módosítása (megabájt, kilobájt, bájt).
- (e) Tartomány másolása.

Megoldás:

(a) feladat:

A fájl importálása a feladat fontos része. A feldolgozandó adathalmaz egy txt kiterjesztésű fájlban van, az adatoknak az adott programba történő importálása elengedhetetlenül fontos a felhasználó számára, hiszen ennek sikertelensége esetén nem tudja végrehajtani a feladatot. Mielőtt elkezdenénk a feladat megoldását, célszerű az adatokat egy szövegszerkesztő alkalmazással megnyitni és megtekinteni. Válasszuk ehhez a MS Wordöt. A következőt láthatjuk:

```
Thu Nov 14 09:00:02 MET 1996
IP accounting rules
pkts bytes prot source      destination    ports
8518 531K all anywhere    157.181.76.0/24 n/a
12887 540K all 157.181.76.0/24 anywhere      n/a
Thu Nov 14 10:00:02 MET 1996
IP accounting rules
pkts bytes prot source      destination    ports
5479 514K all anywhere    157.181.76.0/24 n/a
4566 261K all 157.181.76.0/24 anywhere      n/a
```

Az első sor jelenti a mérés dátumát. A második sor szövege a feladat célja szempontjából nem tartalmaz információt. A harmadik sor első két szava („pkts” és „bytes”) mértékegységeket jelöl; a következő, a negyedik sor a harmadikban megadott mértékegységek szerint a hálózaton küldött csomagok adatait tartalmazza. Az első adat 8518. Ez ebben a feladatban (és általános esetben is) lehet szám vagy szöveg (jelen esetben szám). A második adat 531K, szintén lehet szám vagy szöveg (jelen esetben szöveg). Az ötödik sor felépítést tekintve ugyanaz, mint a negyedik, a különbség mindössze annyi, hogy a kimenő csomagok és bájtok mennyiségét tartalmazza.

Lépünk be a „Fájl” menüpontba! A „Megnyitás” almenüpontban lépünk be és keressük meg az ACCT.txt nevű fájlt. Ha nem látjuk, a fájlnev címszóban állítsuk át a „Minden Excel fájltypus”-t „Minden fájl”-ra. Válasszuk ki az ACCT.txt fájlt és nyissuk meg. Ekkor megjelenik a „Szövegbeolvasó varázsló” első lépése. Itt nem mindegy, hogy az adat típusára

milyen választást eszközölünk, ugyanis az alapértelmezés lehet az úgynevezett „Fix széles”. Ekkor, amint a megjelenő ablak alsó részében megtekinthető, egy sor egy cellában jelenik meg, ami azonban kezelhetetlen, ugyanakkor a beolvasandó fájlban valamilyen karakterrel el vannak választva az egyes adatok (mezőhatárok). Ezért válasszuk az adattípusra a „Tagolt”-at és lépünk tovább. Itt több „Határoló jel” közül választhatunk. Azt kell kiválasztani, ami az egyes adatokat (mezőket) jól választja el egymástól, esetünkben a tabulátor megfelelő. Lépünk tovább. Minden oszlop alapértelmezésének az Excel automatikusan az „Általános” típust adja meg. Ez a feladat szempontjából elfogadható, így választhatjuk a „Befejezés” műveletet. A következő képernyőfelületet kapjuk (9. ábra):

A1	:	X	✓	f	Thu		
	A	B	C	D	E	F	G
1	Thu	Nov	14	9:00:02	MET	1996	
2	IP	accounting	rules				
3		pkts	bytes	prot	source	destination	ports
4		8518	531K	all	anywhere	157.181.7	n/a
5	12887	540K	all	157.181.7	anywhere	n/a	
6	Thu	Nov	14	10:00:02	MET	1996	
7	IP	accounting	rules				
8		pkts	bytes	prot	source	destination	ports
9		5479	514K	all	anywhere	157.181.7	n/a
10		4566	261K	all	157.181.7	anywhere	n/a
11	Thu	Nov	14	11:00:02	MET	1996	
12	IP	accounting	rules				
13		pkts	bytes	prot	source	destination	ports
14	16218	2064K	all	anywhere	157.181.7	n/a	
15	22604	966K	all	157.181.7	anywhere	n/a	
16	Thu	Nov	14	12:00:01	MET	1996	
17	IP	accounting	rules				

9. ábra: Az adatimportálás után ezt a képet kapjuk

Az elvégzett műveletek után elkezdhető a feladat megoldása.

(b) feladat:

Az első lépésben a fenti táblázatot át kell alakítani olyan módon, hogy az egymás alatt lévő (egy oszlopba kerülő) adatok ugyanannak a paraméternek mérési értékeit tartalmazzák. Ennek célszerű formája a következő lehet:

Oszlop jele	I	J	K	L	M
Paraméter neve	Dátum, idő	Bejövő csomag	Bejövő bájtt	Kimenő csomag	Kimenő bájtt
Példa 1	1996.11.14 9:00:02	8518	543744	12887	552960
Példa 2	1996.11.14 10:00:02	5479	526336	4566	267264

Az egyszerűség kedvéért az új táblázatot az I oszloptól kezdve helyezzük el.

Első lépésként a dátumot és az időt célszerű „helyre rakni”. Ehhez használjuk a DÁTUMÉRTÉK(C1&"-"&B1&"-"&F1)+D1 függvényt az előző fejezetben bemutatottak szerint. A lehetséges bemenő formátumok közül a következőt választottuk: „Nap-Hó-Év”.

I1 → =DÁTUMÉRTÉK(C1&"-"&B1&"-"&F1)+D1

Ezzel a dátumot várakozásainknak megfelelő formára hoztuk. Az I1 cellában megjeleníteni kívánt dátum helyett annak csupán egy része jelent meg. A teljes dátum megjelenítése érdekében a menü segítségével végezzük el a következő műveleteket: az I1 cellában állva a jobb egérgombbal előhívható legördülő menü „Cellaformázás” opciójának segítségével válasszuk ki az "Egyéni" lehetőséget, és a „Formátumkódba” írjuk be a következőket: „éééé.hh.nn óó:pp:mm”.

Ezt követően az I1 cellában sok #-et láthatunk. Növeljük meg az oszlop szélességét az egér segítségével, így az I1 cella tartalma és értéke már az elvárásainknak megfelelő lesz.

(c) feladat:

Nagyobb problémát jelent a J-M oszlopok kialakítása, hiszen az Excel nem teljesen olyan formában olvasta be az adatokat, mint elvártuk, hiszen pl. a 4. sor a többihez képest egy oszloppal elcsúszott jobbra. Ha a táblázatunk nem lenne ennyire hosszú, ezeket a hibákat kézzel kijavíthatnánk, de a többszáz- vagy ezersoros táblát soronként javítani nagyon nehéz. Ezért „automatizálni” kell a javítást. Szerencsére elég egyszerű dolgunk van, hiszen megállapítható, ha a rekord 4. sorának első oszlopában van tartalom, akkor nincs elcsúszva a sor, ha viszont ez a cella üres, akkor el van csúszva. Célszerű egy külön oszlopban elvégezni a javítást; erre használjuk a N-Q oszlopokat. A javítás „algoritmusa” a következő lehet:

- Ha a rekord 4. sorának első cellája üres, akkor a rekord elcsúszott,
- de ha nem üres, akkor nem csúszott el.

Ugyanezt alkalmazzuk majd az 5. sorra. A J és K oszlopba a 4. sorból, az L és M oszlopba az 5. sorból fognak átszármazni az adatok. A J-M oszlopok helyet használjuk az N-Q oszlopokat (később meglátjuk, miért). A fenti „algoritmus” egy feltételt tartalmaz: a feltételes kifejezésekhez a már ismert HA(...) függvényt lehet használni. Esetünkben a legegyszerűbb a következő alakban:

N1 → =HA(\$A4="" ;B4;A4)

majd ezt a kifejezést másoljuk át az O1 cellába. A P1 cella ehhez teljesen hasonló lesz, csak a 4. sor helyett az 5. sort kell használni, így:

P1 → =HA(\$A5="" ;B5;A5)

és ezt másoljuk át a Q1 cellába. Ezzel a N-Q oszlopok tartalma már megfelelő.

(d) feladat:

A rendezés után az O1 és Q1 cellák még tartalmazzák a "K" és "M" jeleket, hiszen ezeket a cellákat az Excel nem numerikus, hanem szöveges cellaként érzékeli.

A következő lépésben ezeket kell feldolgozni. Erre a következő „algoritmust” használhatjuk:

- Ha a O1:Q1 cellák tartalmának utolsó betűje „K”, akkor a tartalomtól azt elhagyva, a maradékot számként kezelve meg kell szorozni 1024-gyel.
- Ha O1:Q1 cellák tartalmának utolsó betűje „M”-re végződik, az előző esethez hasonlóan kell eljárni, de nem 1024-el, hanem 1024x1024-gyel kell szorozni.
- Ha sem „K”-val, sem „M”-el nem végződik, akkor változatlanul kell hagyni.

A fenti feladatokat egyetlen összetett HA függvény segítségével is meg lehet oldani, amelyet a következőképpen építünk fel:

1. Az adott cella (az első esetben az N1-nek megfelelő J1) utolsó karakterét a JOBB(N1;1) függvénnyel kapjuk meg.
2. A cellatartalom – az utolsó karakter nélkül – a BAL(N1;HOSSZ(N1)-1) kifejezéssel kapható meg.
3. A 2. pontban kapott mennyiség numerikus értékévé az ÉRTÉK(..) függvény segítségével alakítható.

A fenti ismétlődő feladatok után kezdjük el felépíteni az utasítást (először csak a „K”-s alakkal foglalkozzunk):

=HA(JOBB(N1;1)="K";1024*ÉRTÉK(BAL(N1;HOSSZ(N1)-1));N1)

Teljesen hasonló lesz az „M”-ek kifejezése:

=HA(JOBB(N1;1)="M";1024*1024*ÉRTÉK(BAL(N1;HOSSZ(N1)-1));N1)

Ezt a kifejezést az előző HA függvény legvégére, az N1 helyére kell beírni, hiszen azt a kifejezést fogja az Excel kiértékelni, ha a cella utolsó karaktere nem „K”. Egymásba ágyazott HA függvények esetén mindig a hamis kimenetbe kell beágyazni az új függvényt. Így a J1 cellába kerülő teljes kifejezés a következő lesz:

J1→=HA(JOBB(N1;1)="K";1024*ÉRTÉK(BAL(N1;HOSSZ(N1)-1));HA(JOBB(N1;1)="M";1024*1024*ÉRTÉK(BAL(N1;HOSSZ(N1)-1));N1))

Ez után másoljuk át az J1 cellát az J1:M1 tartományba, hiszen a többi oszloppal is ugyanezt kell tennünk **(10. ábra)** (bármelyik oszlopban előfordulhat, hogy a cella tartalma K, M vagy szám végű).

[illegible]

(e) feladat:

Ezzel az első óra (nevezhetjük rekordnak is) feldolgozásával végeztünk. Ahhoz, hogy az összes rekordot így alakítsuk át, már csak az I1:Q1 cellákat kell átmásolni minden 5. sorba, mivel a program 5 soros blokkokban rögzítette az adatainkat. A létrehozott struktúrát úgy kell elképzelni, mintha „lefotóznánk” az első öt sort (I1:Q5) – beleértve az adatkinyerésre használt celláinkat – és azt önmaga alá másolnánk annyiszor, ahányszor előfordul a szekvencia. A lényeg, hogy a kijelölt tartomány sorainak száma egész számú többszörösét jelöljük ki. Erre az alábbi lehetőséget kínálja az Excel: jelöljük ki az I1:Q5 tartományt, és adjuk ki a „másolás” (CTRL+C) parancsot. Céltérületnek jelöljük ki (a SHIFT folyamatos lenyomva tartása mellett) I1-től azt a tartományt, amíg az adataink tartanak (I1:Q500) (szükség esetén a PAGE DOWN vagy függőleges nyíl – kurzormozgató – billentyűket), majd üssük le az ENTER billentyűt. Ha nem sikerült $n \cdot 5$ sort kijelölni, akkor az Excel egy hibaüzenetet küld, ekkor jelöljük ki több vagy kevesebb sort, és ezt egészen addig ismételjük, amíg az Excel el nem végezi a másolást (ezt legfeljebb 4 kísérletből meg kell tudnunk tenni) **(11. ábra)**.

[illegible]

Ezzel létrejött egy nekünk megfelelő táblázat, azzal a „hibával” hogy sok üres sor szerepel benne. Ezeket a legegyszerűbben úgy tudjuk eltávolítani, hogy CTRL+C paranccsal vágólapra másoljuk a I1:Qn (n a táblázatunk sorainak száma, esetünkben 500) tartományt, és az I1 cellába bemásoljuk az adatokat a „Jobb egérgomb/Írányított beillesztés/Értéket” parancs segítségével. Ezzel megszüntetjük a képleteket, azaz csak azok értékét őriztetjük meg, majd az idő szerint sorba kell rendeznünk az adatokat. Ezt követően kijelöljük az A1-H1 cellákat és a „Törlés/Egész oszlop” menü segítségével kitöröljük a felesleges oszlopokat. Ugyanígy kitöröljük az F-I oszlopokat is (hiszen ezek „mellékszámítások” voltak).

Végezetül készítsünk fejléct a táblázathoz: álljunk az A1 cellára (a CTRL+HOME billentyűkombináció lenyomásával), és a „Beszúrás/Egész sor” helyi menü segítségével szúrjunk be egy üres sort a táblázat elejére. Írjuk be a következőket:

A1 → Idő

B1 → Bejövő csomagok

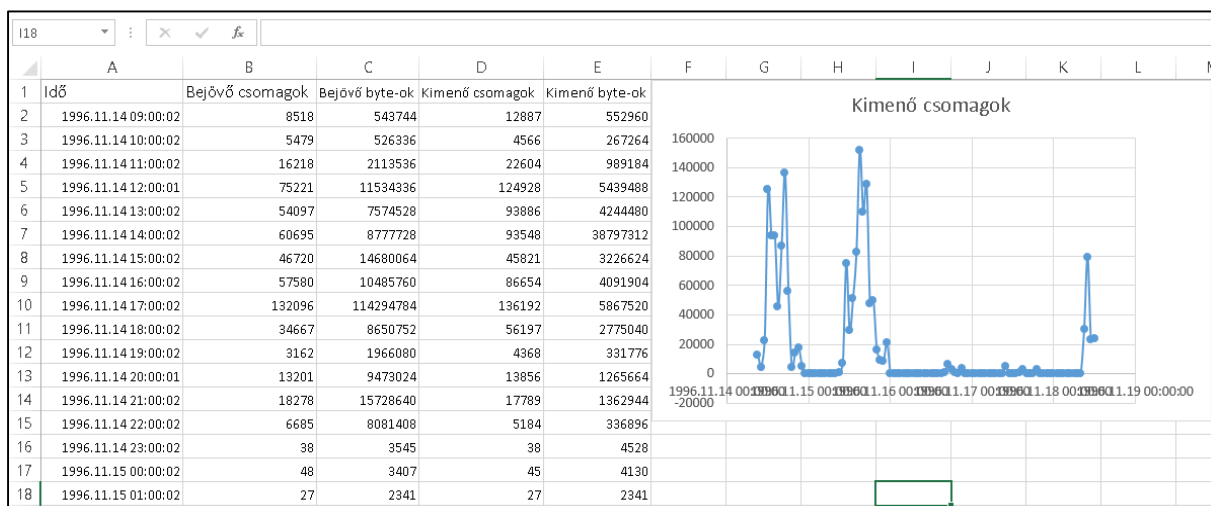
C1 → Bejövő bájtok

D1 → Kimenő csomagok

E1 → Kimenő bájtok

Ezzel a táblázatunkat olyan formátumúvá alakítottuk, amely lehetővé teszi a további számítások elvégzését vagy ábra készítését.

Példaként ábrázoljuk grafikonon a „Kimenő csomagok” mennyiségét az idő függvényében: jelöljük ki az A és D oszlopokat, majd a „Beszúrás/Pont(x,y)” ábra segítségével készítsük el az alábbi grafikont (12. ábra).



12. ábra: Az (e) feladat megoldása

5. Az FKERES függvény használata felszíni vízminőségi adatok esetén

Példafájl neve: fkeres_feszini_vizmin_adatok.xlsx

Adatforrás: nem nyilvános

A feladat célja:

Számos esetben előfordul, hogy egy hosszú távú monitoringrendszer működtetése során az egyes paramétereket külön-külön gyűjtik táblázatokba (pl. paramétercsoportonként). Ez jelentős problémát okozhat abban az esetben, ha például sokváltozós adatelemzést szeretnénk végezni. Ezért gyakori feladat egy egységes, az összes paraméter adatait tartalmazó táblázat létrehozása. A feladat során tehát külön munkalapokon található felszíni vízminőségi adatsorok egységes táblázatba rendezését mutatjuk be az FKERES függvény segítségével.

Új ismeretek:

- FKERES függvény használata
- tartományok, adattábla elnevezése
- munkalapok közötti hivatkozások

Elméleti háttér (TATÁR ÉS ÓVÁRI EDS., 2012 alapján):

A vízminőség védelme kiemelt szereppel bír a környezettudományokban, hiszen a víz a legjelentősebb transzportközeg az élőlények számára, emellett élőhely is egyben. Ezért kiemelten fontos feladat vizeink megfelelő mintázása, ezen keresztül kimerítő ismerete (KOVÁCS ET AL., 2012c). A vízmintavétel során számos olyan változót meghatározunk, amelyek elengedhetetlenek vizeink átfogó minősítéséhez. Ezeket több csoportba sorolhatjuk. A legegyszerűbb megközelítés szerint beszélhetünk fizikai (pl. vízhozam), kémiai (pl. különböző ionok koncentrációi) vagy biológiai (pl. klorofill-a koncentráció) paramétereiről. Egy másik megközelítés során külön csoportba sorolhatjuk a felépítő szervezetek számára fontos tápanyagokat (pl. különböző N és P formákat). Mintavételezés során alapvető fontosságú rögzíteni a mintavétel helyét és idejét is.

Feladatban felhasznált függvények, funkciók:

- **FKERES függvény:**

Az FKERES függvény egy adattábla első oszlopában keres egy megadott értéket, majd eredményül megadja a találatnak megfelelő sorban lévő, meghatározott cella értékét

=FKERES(keresési_érték; tábla; oszlop_száma; [tartományban_keres])

keresési_érték: Az adattábla első oszlopában keresendő érték. Ha ez kisebb, mint a tábla első oszlopának legkisebb értéke, akkor a függvény a „#HIÁNYZIK” hibaértéket adja

eredményül. Ha a keresési érték nagyobb, mint az adattábla első oszlopának legnagyobb eleme, akkor viszont az alapértelmezés szerint a tábla legnagyobb eleméhez tartozó értéket fogja eredményül adni. (Ennek okát később, a *[tartományban_keres]* argumentum bemutatása során részletezzük.)

tábla: Az adatokat tartalmazó cellatartomány. A tábla argumentum lehet tartomány (például A2:D8), vagy egy tartomány neve. A függvény a kis- és nagybetűk között nem tesz különbséget. A *tábla* argumentumban rögzített cellatartomány első oszlopában lévő értékek között keresi a függvény a keresési értéket.

oszlop_szám: A tábla azon oszlopának a táblán belüli sorszáma, amelyből a *keresési_érték*nek megfelelő értéket kiválasztja a függvény. Amennyiben ez az érték egynél kisebb, az eredmény „#ÉRTÉK!” lesz; abban az esetben, ha az *oszlop_szám* értéke nagyobb, mint a tábla oszlopainak száma, akkor a „#HIV!” hibaértéket adja eredményül a függvény.

[tartományban_keres]: Logikai érték, amellyel a függvényhez pontos vagy közelítő keresést írhatunk elő. Ezen argumentum megadása nem kötelező. Amennyiben nem adjuk meg, akkor a függvény alapértelmezésben IGAZ beállítással keres.

Ha a *tartományban_keres* értéke IGAZ (vagy hiányzik), akkor az FKERES által visszaadott érték könnyen megtévesztő lehet. Abban az esetben, ha a függvény nem talál pontos egyezést a tábla első oszlopában, akkor a következő legnagyobb, de a *keresési_érték*nél kisebb érték sorából választ ki egy értéket (az oszlop számnak megfelelően).

Ha a *tartományban_keres* argumentum értéke HAMIS, akkor az FKERES csak pontosan egyező értéket keres. Ha a tábla első oszlopában két vagy több, a *keresési_érték* argumentummal egyező érték szerepel, akkor a függvény az első értéket használja. Ha pontos egyezés nincs, akkor a „#HIÁNYZIK” hibaértéket adja eredményül.

Feladat:

Gyakran előfordul, hogy egy mintavételezési kampány során egyszerre begyűjtött vízmintákból létrejövő adatokat, paraméterenként külön-külön munkalapra/táblázatba rendezik. Tegyük fel, hogy egy laboratórium az egyes felszíni vízminták adatait három különböző táblázatba gyűjtötte, a paraméterek típusa szerint csoportosítva. Az első munkalap tartalmazza a vízhozam és pH változó értékeit, a második az ionokat, végül a harmadik a tápanyag- és oxigénháztartás paramétereinek értékeit. Minden munkalapon 1990.01.08-tól 1991.12.17-ig szerepelnek az adatok. A vízhozamot és a pH-t minden esetben meghatározták, ezért ez a táblázat tartalmazza az összes mérési időpontot. De a további két paramétertípus esetében a sorok száma eltérő, mert volt olyan mintavételezés, amikor vagy az ionokat vagy a tápanyagokat nem mérték meg.

- (a) A feladat: összevonni a három táblázatot egyetlen táblázatba, ahol az egyes értékek a megfelelő időponthoz tartoznak, még hozzá úgy, hogy minden időpont csak egyszer szerepel! (Azokban az időpontokban, amikor nincs például ion adat, szerepeljen üres cella a megfelelő paramétereknél).

Megoldás:

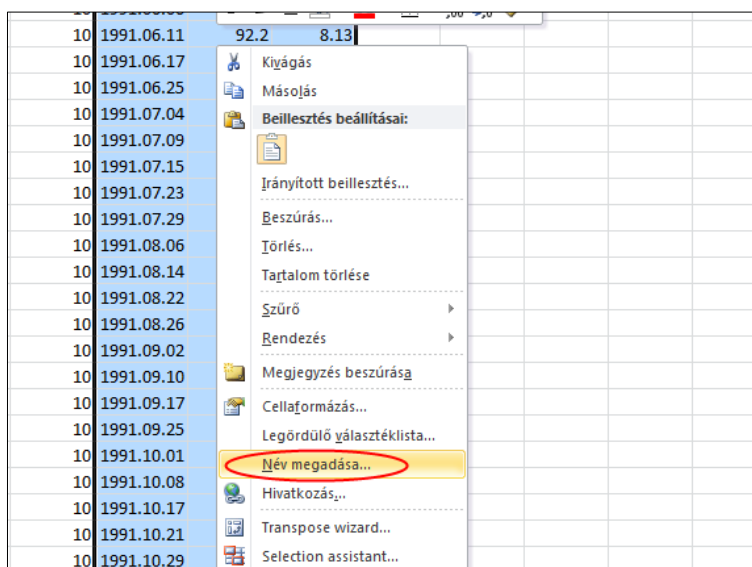
Első lépésként vizsgáljuk meg, hogy milyen adatok állnak rendelkezésre! Ehhez nyissuk meg a példafájlt (*fkerez_feszini_vizmin_adatok.xlsx*).

A fájl három munkalapot tartalmaz a paraméterekkel. Ellenőrizzük, hogy az egyes munkalapokon lévő paraméteroszlopok hossza valóban eltérő-e.

Hozzunk létre egy új munkalapot „Megoldás” névvel. Mivel a pH-t és a vízhozamot minden méréskor meghatározták, így a hozzájuk tartozó időtengely bizonyosan tartalmazza az összes mérési időpontot, ezért másoljuk át a „vízhozam+pH” munkalap első 3 oszlopát a „Megoldás” munkalap A1 cellájába! Ezuán készítsük el a „Megoldás” munkalapon az összerendezett táblázat fejlécét, hogy a megfelelő oszlopokba rendezhessük az összes paramétert! Másoljuk át a D1:R1 cellákba a paraméterek neveit megfelelő sorrendben (a „Vízhozam+pH” munkalap D1:E1, az „ionok” munkalap D1:J1, és a „tápanyag” munkalap D1:I1 celláit).

Ezután határozzuk meg az FKERES függvényhez szükséges *táblákat* (ezek pedig az adatokat tartalmazó cellatartományok). Összesen három táblát kell definiálnunk, a háromféle paramétertípus szerint.

- Jelöljük ki a „vízhozam+pH” munkalapon a C2:E105 cellatartományt! Mivel a dátumok alapján szeretnénk a vízhozamot, valamint a pH-t megtalálni, szükség van a C oszlopra is. Jobb egérgombbal kattintsunk a kijelölt területre, majd válasszuk a „Név megadása” opciót (**13. ábra**). A megjelenő ablakban a „Név tartomány” kitöltésével nevezzük el a cellatartományt „pH” névvel. Később az FKERES függvényben majd „pH” néven tudunk erre a táblára hivatkozni.



13. ábra: Tartomány elnevezése Excelben

- Jelöljük ki az „ionok” munkalapon a C2:J78 cellatartományt, majd a fent leírt módon nevezzük el a tartományt „ionok” néven!
- Végül a „tápanyag” munkalapon lévő C2:I83 tartományt nevezzük el „tápanyag” néven.

Annak érdekében, hogy a műveletet valamelyest automatizáljuk, számozzuk be a paraméterek oszlopait annak függvényében, hogy az adott paraméter hányadik oszlopban található a saját

névvel ellátott táblájában. Ehhez szúrjunk be egy sort a „Megoldás” munkalapon, a táblázat fölé. A D1:R1 tartományba a következő számsor kerüljön: 2, 3, 2, 3, 4, 5, 6, 7, 8, 2, 3, 4, 5, 6, 7.

Magyarázatul: a D1 cellába például azért került 2, mert a D oszlopnak megfelelő „Vízhozam” oszlop, a „Vízhozam+pH” munkalap, „pH” elnevezésű adattáblájának második oszlopában található. Az R1 cellában azért kerül 7 mert a „tápanyag” munkalap „tápanyag” nevű tartományának a 7. oszlopában található a foszfát-foszfor paraméter (PO₄-P).

Kezdjük el feltölteni a „Megoldás” munkalap táblázatát a paraméterek értékeivel! Mivel három különböző táblából keressük ki a paraméterek adatait, ezeket külön kell kezelnünk. Először keressük ki a „Vízhozam” Ca és oldO paraméterek megfelelő értékeit FKERES függvényvel:

D3 → =FKERES(\$C3;pH;D\$1;HAMIS)

F3 → =FKERES(\$C3;ionok;F\$1;HAMIS)

M3 → =FKERES(\$C3;tápanyag;M\$1;HAMIS)

Ezután másoljuk át D3 tartalmát az E3 cellába, F3 tartalmát a G3:L3 tartományba, illetve M3 tartalmát N3:R3 tartományba! Végül másoljuk le a teljes D3:R3 sort a táblázat aljáig, tehát a 106. sorig!

Szélesítsük meg az oszlopokat, ahol szükséges, majd végezetül szüntessük meg az egyes cellák „#HIÁNYZIK” tartalmát! Azokban a cellákban szerepel „#HIÁNYZIK” tartalom, ahol az adott paraméter nem volt mérve, azaz ezek a cellák valójában üres cellák. Jelöljük ki a D3:R106 tartományt, majd másoljuk a vágólapra, végül illesszük be értékként ugyanarra a területre! (A terület kijelölése után CTRL+C paranccsal másoljuk a vágólapra a táblázatot, majd jobb kattintás után, a legördülő menüben a „Beillesztés beállításai” menüpont alatt válasszuk az „Értékek” opciót!) Így a cellákban eddig szereplő képleteket azok értékével helyettesíthetjük. Ezután jelöljük ki a D3:R106 tartományt újra, és CTRL+H billentyűkombinációval nyissuk meg a „Keresés és csere” ablakot. Itt a „Keresett szöveg” helyére írjuk „#HIÁNYZIK” szöveget, a „Csere erre” helyét hagyjuk üresen! Kattintsunk az „Összes cseréje” gombra, így végül törölhetjük a segédszámozást tartalmazó első sort.

Összefoglalva: bemutattuk az FKERES függvény működését, elneveztünk tetszőleges tartományokat későbbi hivatkozhatóság céljából, illetve bemutattuk azt, hogy hogyan lehet különböző munkalapok tartalmát összerendezni egy keresési funkcióval.

6. Az FKERES és VKERES függvény alkalmazása harmatpont meghatározására, meteorológiai adatokból

Példafájl neve: fkeres_vkeres_harmatpont.xlsx

Adatforrás: Weather Underground, <http://www.wunderground.com/>

A feladat célja:

Az FKERES és a VKERES függvény logikai összekapcsolása és alkalmazása összetett táblázatban történő kereséshez. Meteorológiai adatok esetén gyakran előfordul, hogy egy számítandó paraméter értékét két másik paraméter értékének együttes változása befolyásolja. Ilyen esetekben egy segéd-adatmátrix segítségével kereshetjük ki a minket érdeklő paraméter értékét. Másik lehetőség, hogy a paraméter meghatározásához használható képletet készítjük el az Excel munkafüzetben. Ebben a feladatban meteorológiai adatokból (léghőmérséklet, relatív páratartalom) fogjuk a levegő harmatpontját meghatározni.

Új ismeretek:

- FKERES és VKERES függvény összekapcsolása
- Egy matematikai képlet bevitele Excelbe
- LOGARITMUS függvény
- HATVÁNY függvény
- MINIMUM és MAXIMUM függvény

Elméleti háttér (<http://www.srh.noaa.gov/>):

A felhőképződés meghatározásához nagyon fontos a harmatpont ismerete. A harmatpont az a levegőhőmérséklet, amelyen a levegő relatív páratartalma 100% lesz. A levegőt tovább hűtve elkezdődik a pára kicsapódása. A felhőképződés során a következő folyamat játszódik le: a meleg levegő passzív felemelkedése és kitágulása hűlést idéz elő, amely így a harmatpontot elérve túltelítetté válik, és kondenzációs magokon megindul a vízcseppek keletkezése. A harmatpont meghatározásához, használhatunk táblázatot, amelyben a relatív páratartalom és a léghőmérséklet függvényében vannak megadva a harmatpontértékek, illetve kiszámolhatjuk a következő képletek segítségével:

$$e_s = 6,11 \times 10^{\left(\frac{7,5 \times T}{237,3 + T}\right)}$$

amelyben, az e_s a telített gőznyomás, a T pedig a levegő aktuális hőmérséklete (°C)

$$T_d = \frac{237,3 \times \ln\left(\frac{e_s \times r_h}{611}\right)}{7,5 \ln 10 - \ln\left(\frac{e_s \times r_h}{611}\right)}$$

amelyben a T_d a harmatpont hőmérséklet (°C), az r_h a relatív páratartalom (%)

Feladatban felhasznált függvények:

- **FKERES és VKERES függvény:** ismétlésképpen, a két függvény felépítése a következő:

„=FKERES(keresési_érték; tábla; oszlop_száma; [tartományban_keres])”

„=VKERES(keresési_érték; tábla; sor_száma; [tartományban_keres])”

- **LOG és LN függvény:** LOG függvény egy adott szám tetszőleges alapú, míg az LN függvény egy adott szám természetes alapú logaritmusát számítja ki.

„=LOG(száma; alap)”

„=LN(száma)”

- **HATVÁNY függvény:** egy szám tetszőleges kitevőjű hatványát számolja ki.

„=HATVÁNY(száma; kitevő)”

- **MIN és MAX függvény:** megadja a bemeneti adatok (számok vagy tartomány) minimum illetve maximum értékét.

„=MIN(száma1; száma2; ... vagy tartomány)”

„=MAX(száma1; száma2; ... vagy tartomány)”

alap: tetszőleges logaritmus alap,

kitevő: tetszőleges hatványkitevő,

szám: az a cella amelynek a logaritmusát, minimumát vagy maximumát szeretnénk kiszámolni,

tartomány: azon cellatartomány, amelynek a minimumát vagy maximumát szeretnénk kiszámolni (pl. A2:A100).

Feladatok:

A kapott Excel munkafüzetben az első munkalapon („harmatpont_budapest”) található, a „Budapest, Ferihegy 2” területén található meteorológiai megfigyelőállomás 2001-től 2005-ig terjedő adatsora. Az adatsor napi felbontású, az első oszlopban (A oszlop) a dátum, a második oszlopban (B) a napi középhőmérséklet, míg a harmadik oszlopban (C) a napi átlagos relatív páratartalom-értékek találhatóak.

A feladat első részében a D oszlopban dolgozva határozzuk meg az adott hőmérséklet és a relatív páratartalom adataiból a „tablazat” munkalapon található segéd táblázat felhasználásával a harmatpont értékeket. Ezt követően számítsuk ki a harmatpontot a fejezet első részében leírt képlettel (az E oszlopba kerülnek). Végül az F oszlopba határozzuk meg a két oszlop közötti eltérést, adjuk meg az eltérések minimumát és a maximumát, továbbá ábrázoljuk a harmatpont változását az idő függvényében.

A „tablazat” munkalapon található segéd táblázat felépítése a következő **(14. ábra)**:

C26									-17.7224562159918
	A	B	C	D	E	F	G		
1	Harmatpont (°C)								
2			20.00	25.00	30.00	35.00	40.00		
3		-20.00	-37.09	-34.88	-33.04	-31.46	-30.07		
4		-19.00	-36.24	-34.01	-32.16	-30.56	-29.16		
5		-18.00	-35.39	-33.15	-31.27	-29.66	-28.25		
6		-17.00	-34.54	-32.28	-30.39	-28.77	-27.34		
7		-16.00	-33.70	-31.41	-29.51	-27.87	-26.43		
8		-15.00	-32.85	-30.55	-28.63	-26.98	-25.53		
9		-14.00	-32.00	-29.68	-27.75	-26.08	-24.62		
10		-13.00	-31.16	-28.82	-26.87	-25.19	-23.71		
11		-12.00	-30.31	-27.96	-25.99	-24.29	-22.81		
12		-11.00	-29.47	-27.09	-25.11	-23.40	-21.90		
13		-10.00	-28.63	-26.23	-24.23	-22.51	-20.99		

14. ábra: A segédtáblázat felépítése

A táblázat „B” oszlopában találhatóak a hőmérséklet, míg a második sorában pedig a páratartalom intervallumértékei, az adatmátrixban pedig az ezekhez az értékekhez tartozó harmatpontértékek Celsius fokban kifejezve.

- Határozzuk meg a D oszlopba a táblázat alapján a harmatpontot!
- Határozzuk meg az E oszlopba a képlet alapján a harmatpontot!
- Az F oszlopba határozzuk meg az eltérést a táblázat és a képlet alapján számolt harmatpontértékek között!

Megoldás:

(a) feladat:

Korábban megtanultuk az FKERES és VKERES függvény külön-külön történő alkalmazását. Ebben az esetben azonban nem elegendő csak az egyiket használni, mivel a keresett érték két paraméter értéke alapján határozható meg. A probléma megoldásához olyan módon kell kombinálni a két függvényt, hogy az egyik függvény harmadik argumentumába (kimeneti sor/oszlop) beírjuk a másik függvény a következőképpen:

„=FKERES(keresési_érték1; tábla; VKERES(keresési_érték2; tábla; sor_szám; [tartományban_keres])); [tartományban_keres])”.

Jelen példára alkalmazva, az FKERES függvénnyel megkeressük a hőmérséklet értéket a táblázat első oszlopában, azonban ehhez többféle páratartalom-érték tartozhat, amelyeket a különböző oszlopokban tárolnak. Mivel nem tudjuk megadni a kimeneti oszlop sorszámat, ezért a probléma megoldásához a VKERES függvényt használjuk, amellyel kikeressük a megfelelő értékhez (páratartalom) tartozó kimeneti oszlop sorszámat. Ehhez módosítanunk kell

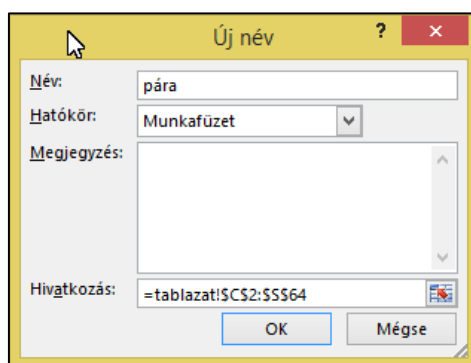
az adatmátrixot, az oszlopokat be kell sorszámozni! Az „A64”-be írjuk be, hogy „Sorszám”, majd „B64”-től sorszámozzuk be az oszlopokat (15. ábra):

61		38.00	11.14	14.55	17.40	19.86	22.04	23.98	25.75
62		39.00	11.95	15.38	18.26	20.74	22.92	24.88	26.66
63		40.00	12.77	16.22	19.11	21.61	23.81	25.78	27.58
64	Sorszám	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00

15. ábra: A segédtábla oszlopainak beszámozása

Ezt követően nevezzük el a tartományokat a „tblázat” munkalapon!

Az FKERES keresési tartománya legyen „hőmérséklet”, B3:S63-ig. Figyeljünk arra, hogy ne vegyük bele a sorszámokat tartalmazó sort! A VKERES keresési tartományának neve legyen „pára”, és C2:S64-ig tartson (16. ábra).



16. ábra: Tartomány elnevezése

Térjünk vissza a „harmatpont_budapest” munkalapra és a D2 cellába írjuk be a függvényt:

D2 → =FKERES(B2;hőmérséklet;VKERES(C2;pára;63))

Az FKERES függvénnel megkeressük a „hőmérséklet” táblában a mért hőmérséklet értékhez (B2) a távozó sort. A VKERES függvény a „pára” tartomány első sorában megkeresi a megfelelő relatív páratartalom-értéket (C2), kimeneti sora az utolsó lesz (63.sor), ahol az oszlopsorszámokat tároltuk. Mivel a segédtáblázatban a hőmérséklet 1 Celsius fokos, a relatív páratartalom pedig 5%-os részletességgel szerepel, ezért közelítő egyezést kell keresnünk. A függvényeknek ez az alapbeállítása, ezért a negyedik argumentumot nem kell megadni.

Ezután az összes sorba másoljuk le a képletet a D1827-es celláig!

(b) feladat:

Ehhez a feladat bevezetőjében található képletet alakítsuk át az Excel formátumára. Elsőként vonjuk össze a két egyenletet:

$$T_d = \frac{237,3 \times \ln \left(\frac{6,11 \times 10^{\left(\frac{7,5 \times T}{237,3 + T} \right)} \times r_h}{611} \right)}{7,5 \ln 10 - \ln \left(\frac{6,11 \times 10^{\left(\frac{7,5 \times T}{237,3 + T} \right)} \times r_h}{611} \right)}$$

Ezek után alakítsuk át a képletet! A hőmérséklet (T) értékei a táblázat B oszlopában, a relatív páratartalom értékei (rh) pedig a C oszlopban találhatóak. Ezek alapján az E2 cellába a következőt írhatjuk:

$$E2 \rightarrow =237.3*LN(6.11*HATVÁNY(10;(7.5*\$B2)/(237.3+\$B2))*\$C2/611)/(7.5*LN(10)-LN(6.11*HATVÁNY(10;(7.5*\$B2)/(237.3+\$B2))*\$C2/611))$$

Figyeljünk a zárójelek megfelelő elhelyezésére, hogy a műveleti sorrend hibátlan legyen! Végül másoljuk le ezt is a táblázat aljáig.

(c) feladat:

Nyilvánvaló, hogy a képlettel számolt érték pontosabb, mivel a segéd táblázat felbontása nem elegendő. Az F2 cellába számoljuk ki az eltérést:

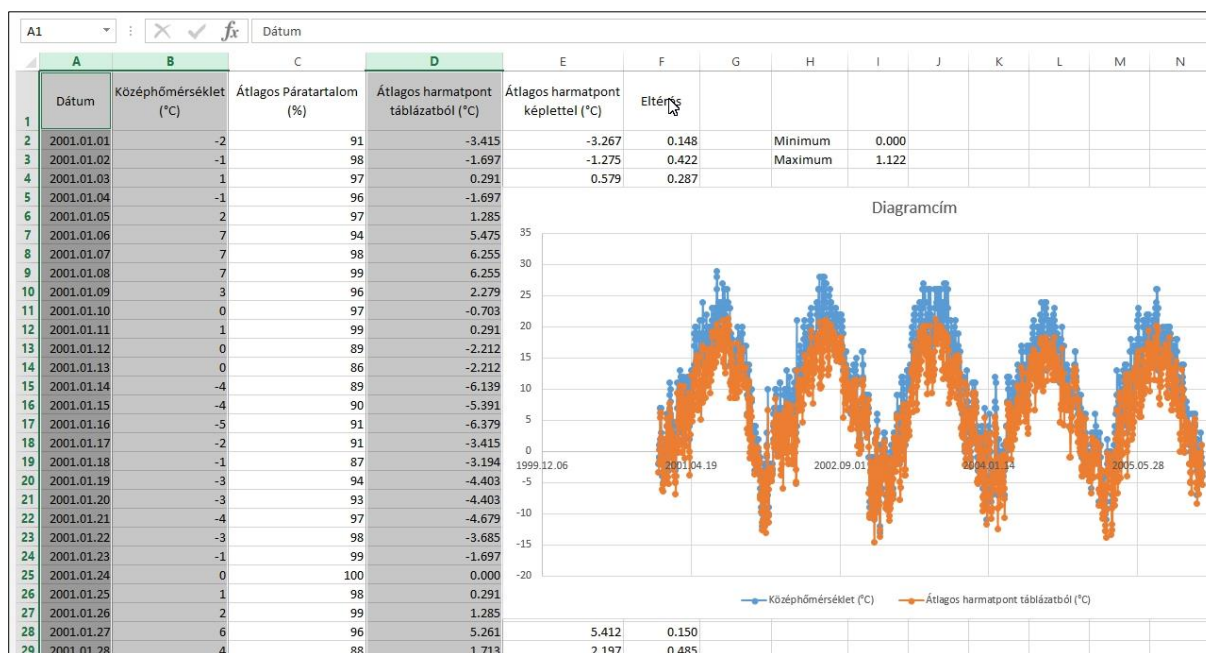
$$F2 \rightarrow =E2-D2$$

Miután az összes sorba lemásoltuk, határozzuk meg az I2 cellába az eltérések minimumát az I3-as cellába pedig a maximumát!

$$I2 \rightarrow =MIN(F2:F1827)$$

$$I3 \rightarrow =MAX(F2:F1827)$$

A legnagyobb eltérés 1,122 °C, amely a táblázat pontatlanságából adódik. Végezetül ábrázoljuk a napi középhőmérsékletet, illetve a képlettel kiszámolt harmatpontértékeket (E oszlop) a dátum függvényében (A oszlop) egy közös pont diagramon (17. ábra).



17. ábra: A napi középhőmérséklet, illetve a harmatpontértékek az idő függvényében

Látható, hogy a harmatpont értéke, akárcsak mint az éves középhőmérséklet, éves periodicitást mutat, a maximuma júniusban, a minimuma decemberben van.

Összefoglalva: megtanultuk az 6. fejezetből, hogy hogyan alkalmazzuk az FKERES és VKERES függvényt adatmátrixban történő keresésre, továbbá azt is, hogy hogyan alakíthatunk át egy bonyolultabb képletet az Excel „nyelvezetére”.

Források:

<http://www.srh.noaa.gov/images/epz/wxcalc/wetBulbTdFromRh.pdf>

<http://elte.prompt.hu/sites/default/files/tananyagok/MeteorologiaiAlapismeretek/index.html>

7. Részösszegképzés a troposzférikus ózonkoncentráció-adatokkal

Példafájl neve: O3.xlsx

Adatforrás: az EMEP³ mérőhálózat interneten elérhető adatbázisa, mely több tíz európai mérőállomásról különböző hosszúságú időintervallumokban gyűjt adatokat.

<http://www.nilu.no/projects/ccc/online/ozone/>

A feladat célja:

A RÉSZÖSSZEG függvény használata, aminek segítségével, gyorsan és hatékonyan lehet létrehozni egy idősor különböző „időfelbontású” összegzéseit: például napi menet óránkénti átlagokból, éves menet havonkénti átlagokból.

Új ismeretek:

- Adatok sorba rendezése
- A RÉSZÖSSZEG függvény használata
- Irányított beillesztés használata

Elméleti háttér (SALMA, 2012 alapján):

A RÉSZÖSSZEG függvény alkalmazását troposzférikus ózonkoncentráció-adatokon fogjuk bemutatni, mivel ezen adatok mérési gyakorisága sűrű, akár órás skálán is történhet, valamint az adatok egyszerre rendelkezhetnek napon belüli és éves ciklikussággal. A XX. század közepéig úgy gondolták, hogy a troposzférikus ózon fő eredete a sztratoszféra, azonban ez a forrás – mint kiderült – csak járulékos a troposzférikus keletkezés mellett. Ilyen troposzférikus keletkezési forrás lehet:

- (i) a villámlás,
- (ii) a növényzet által kibocsájtott szerves vegyületek,
- (iii) a közlekedésből (belső égésű motorokból) származó, elsősorban nitrogén-monoxid oxidációja, majd a keletkezett nitrogén-dioxid fotolízise, vagy
- (iv) a munkahelyi környezetben: főként idősebb fénymásológépek ózonkibocsájtása.

A bemutatott gyakorlatban konkrétan a közlekedésből származó troposzférikus ózonszennyezés esetét fogjuk felhasználni, mivel ezen adatok rendelkeznek mind napszakon belüli (diurnális) ciklikussággal, mind pedig egy karakteres éven belüli menettel is.

Mind a napi, mind pedig az éves ciklikusság kialakulásában jelentős szereppel bír a napsugárzás intenzitása és az UV sugárzás hatására lezajló nitrogén-dioxid fotolízise. Míg az éven belüli ciklikusság a nyáron magasabb számú napsütéses órákhoz köthető, addig a napon-

³ European Monitoring and Evaluation Programme

belüli, a déli/kora délutáni órákra jellemző erősebb napsugárzáshoz (24-32°C; ~70% alatti páratartalom, hőmérsékleti inverzió) és közlekedés megnövekedett intenzitásához köthető.

A jelenség nagyban felelős a Los Angeles típusú (fokokémiai) szmog kialakulásért, melyről részletesebb leírás PONGRÁCZ (2012) és SALMA (2012) tanulmányaiban található, illetve kémiai folyamatairól részletesebben HAAGEN-SMIT 1952-es publikációjában lehet olvasni.

Az ózon élettani hatásai:

- növények levélzetének közvetlen károsítása, égéshez hasonló sérülések
- nyálkahártya és szem irritációja
- légutak funkcionális és morfológiai változásai

Feladatok:

- (a) Ábrázoljuk a rendelkezésre álló troposzférikus ózonkoncentráció-adatok átlagos napi menetét minden 00:00, 01:00, 02:00, 03:00 ... és 23:00 órakor mért ózonkoncentráció átlagaiból mind a hét évre!
- (b) Ábrázoljuk az ózonkoncentráció éves menetét, minden január, február, ... és december havi átlagaiból.

Megoldás:

(a) feladat:

Első lépésként, mint minden gyakorlati példánál, áttekintjük, milyen adathalmazból fogunk dolgozni. Az alábbiakra célszerű figyelni: (i) milyen valószínűségi változó idősorait vizsgáljuk, (ii) milyen sűrűségű volt a mintavétel, (iii) vannak-e hiányzó adatok, és ha igen, akkor azok hogyan jelennek meg, (iv) végül, hogy megfelelően van-e rendezve az adathalmaz.

A feladatfájl (O3.xlsx) betöltése után első lépés, hogy – ha még nincsen – lássuk el fejléccel az adattáblát. (Beszúrunk egy sort az adatok fölé az 1-es sorba, és az A1-es cellába beírjuk, hogy „DÁTUM” a B1-be pedig, hogy „O3 (DU)”, amivel az ózon mértékegységét adjuk meg (Dobson Unit⁴)).

Összesen 61369 sorunk van, ebből egy sor a fejléc. Így tehát 61368 mérési időponttal rendelkezünk, azok viszont nincsenek megfelelően rendezve, hiszen az adatok mért értékük szerint vannak sorba állítva és nem dátum szerint. A dátum szerinti sorba rendezést az „ADATOK” fül „Rendezés” menüpontjával tudjuk elvégezni. Ezzel egyértelművé válik majd, hogy óránkénti mérésekkel van dolgunk. Jelöljük ki a teljes adattáblát a CTRL+A billentyűkombináció használatával, majd kattintsunk az „ADATOK” fül „Rendezés” menüpontjára. Itt fontos bejelölni, hogy az adatok tartalmazznak-e fejléctet (a mi esetünkben igen), majd kiválasztjuk, hogy melyik változó („Oszlop”) milyen tulajdonsága („Rendezés

⁴ Az ózonréteg oszlopsűrűségének mértékegysége. Ha egy adott alapterületű levegőoszlopban lévő összes ózont a Föld felszínén egyenletesen szétoszlátánánk, 1 DU-nak megfelelő mennyiség 1 atmoszféra légnyomáson, 0°C hőmérsékleten 0,01 mm vastag réteget képezne (PAC, 1990).

alapja”) mentén és milyen sorrendbe („Sorrend”) rendezzük az adatokat a következő mondat szerint:

A rendezést a „DÁTUM” – „Értékek” és „Legrégebbi legújabbig” alapján végezzük el. Ezzel időrendi sorrendbe rendeztük az ózonkoncentráció-adatokat. Amennyiben nincsen fejléc, az adott oszlop legördülő menüben az oszlop betűjele jelenik meg.

Az adatok rendezése után látszik, hogy a mérések 1990.01.01-től 1996 végéig történtek óránkénti felbontással, figyelembe véve a szökőéveket, melyekből kettő esik a vizsgált időtartományba (1992 és 1996). Így jön ki a 61368 adat, $(7 \text{ év} * 365 \text{ nap} + 2) * 24 \text{ óra}$.

Az adathiányokat helyesen jelölték: ott ahol nem volt mérés, üres cella van az adatsorban. Elfogadható megoldás lenne, ha az adathiányt egy „irreális” értékkel pl. -999-el, 999-el, vagy „NA”-val jelölnénk. A lehető legrosszabb megoldás, ha 0 értéket írnak azon cellákba ahol nem volt mérés, mivel a kutató adott esetben nem tudja eldönteni, hogy az adott időpontban valóban nem mértek, vagy a valószínűségi változó értéke volt nulla.

Első lépésként létrehozuk a csoportosító változókat, ami alapján majd az átlagot számolni fogjuk:

C1 → ÓRA

D1 → HÓNAP

Következő lépésként „kiszedjük” az A oszlopban lévő dátumból az órát és hónapot a C és D oszlopokba:

C2 → =ÓRA(A2)

D2 → =HÓNAP(A2)

Majd lemásoljuk a C2:D2 blokkot a táblázat aljáig, a 61368. sorig.

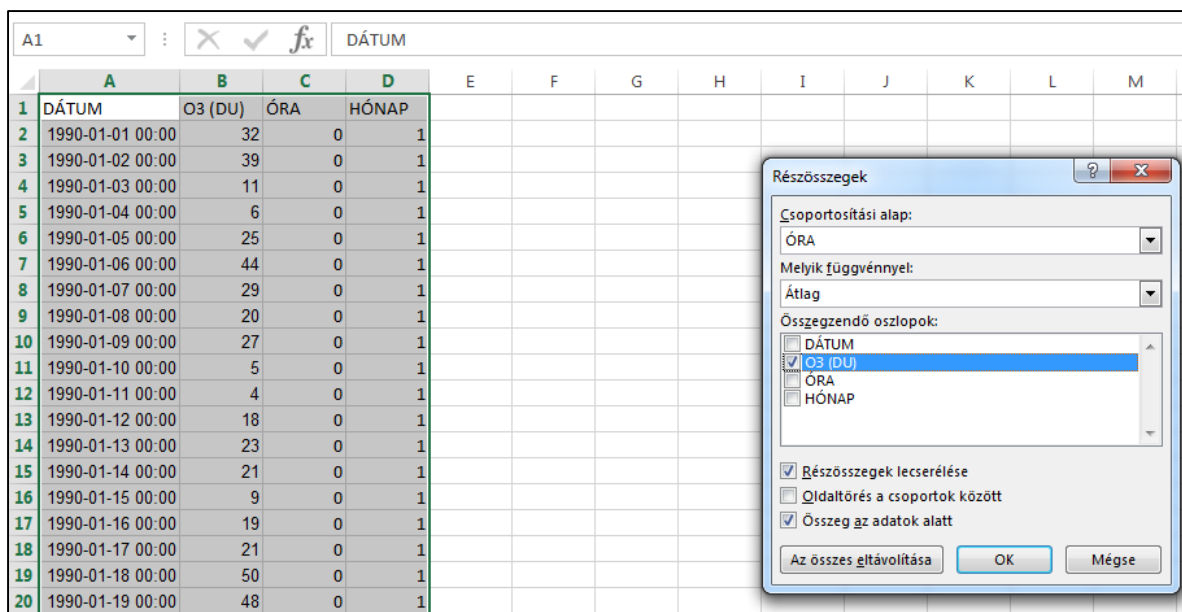
Létrehoztuk azon oszlopokat, amelyek a részösszegképzés (csoportosítás) alapját adják, így elkészíthetjük első körben az órák alapján történő átlagolást.

Részösszegképzés esetén három tényezővel kell számolnunk: (i) a csoportosítás alapja, (esetünkben először az ÓRA), (ii) melyik függvénnyel kívánunk csoportosítani (esetünkben az ÁTLAG) és végül (iii) a csoportosítandó változó (jelen esetben O3 (DU)).

Részösszegképzésnél kiemelten fontos, hogy mindig a csoportosítás alapját adó változó szerint legyen sorba rendezve az adathalmazunk, első esetben az ÓRA szerint, majd a havi átlagok képzésénél a HÓNAP szerint.

Az előzőekhez hasonlóan kijelöljük a teljes adattáblát és sorba rendezzük az ÓRA szerint, majd elkezdjük a részösszegképzést az „ADATOK” menü „Részösszeg” parancsával.

A megnyílt „Részösszegek” párbeszédpanelen kiválasztjuk az alábbiakat az óránkénti átlagok képzéséhez **(18. ábra)**.



18. ábra: A Részösszegeképzés párbeszédpanele

Csoportosítási alap: ÓRA

Melyik Függvénnel: Átlag

Összegzendő oszlopok: O3 (DU)

Az Excel a munkalapot a létrehozás során használt feltételek alapján tagolja. A kapott részösszeztáblát úgy lehet elképzelni, mint egy „harmonikát”, amelyet az ablak bal felső sarkában lévő számozott „szintgombokkal” [1, 2, 3] bont ki és hajt össze a felhasználó. Az [1]-es gomb esetén csak a teljes adattábla összesített részösszeget látjuk ez a főösszeg az 58,95 (a teljes O3 oszlop átlaga) (19. ábra).

F7					
1	2	3	A	B	C
	1	DÁTUM	O3	ÓRA	HÓNAP
+	61394	59 Teljes átlag			
	61395				

19. ábra: A teljes adattábla összesített részösszege [1]

A [2]-es esetén az óránkénti átlagokat és a teljes átlagot látjuk, ezek az egyes csoportok részösszegei (20. ábra).

F7		:	X	✓	f _x	
1	2	3	A	B	C	D
	1	DÁTUM	O3	ÓRA	HÓNAP	
+	2559		47.51556	0 Átlag		
+	5117		45.66007	1 Átlag		
+	7675		44.03104	2 Átlag		
+	10233		43.06492	3 Átlag		
+	12791		41.90096	4 Átlag		
+	15349		41.11893	5 Átlag		
+	17907		42.19331	6 Átlag		
+	20465		45.75028	7 Átlag		
+	23023		51.26432	8 Átlag		
+	25581		58.22993	9 Átlag		
+	28139		65.6124	10 Átlag		
+	30697		72.0889	11 Átlag		
+	33255		76.99594	12 Átlag		
+	35813		80.28474	13 Átlag		
+	38371		81.63681	14 Átlag		
+	40929		81.02283	15 Átlag		
+	43487		79.13488	16 Átlag		
+	46045		74.32389	17 Átlag		
+	48603		68.56926	18 Átlag		
+	51161		62.68276	19 Átlag		
+	53719		57.84894	20 Átlag		
+	56277		53.83058	21 Átlag		
+	58835		51.20451	22 Átlag		
+	61393		49.01533	23 Átlag		
-	61394		59	Teljes átlag		
	61395					

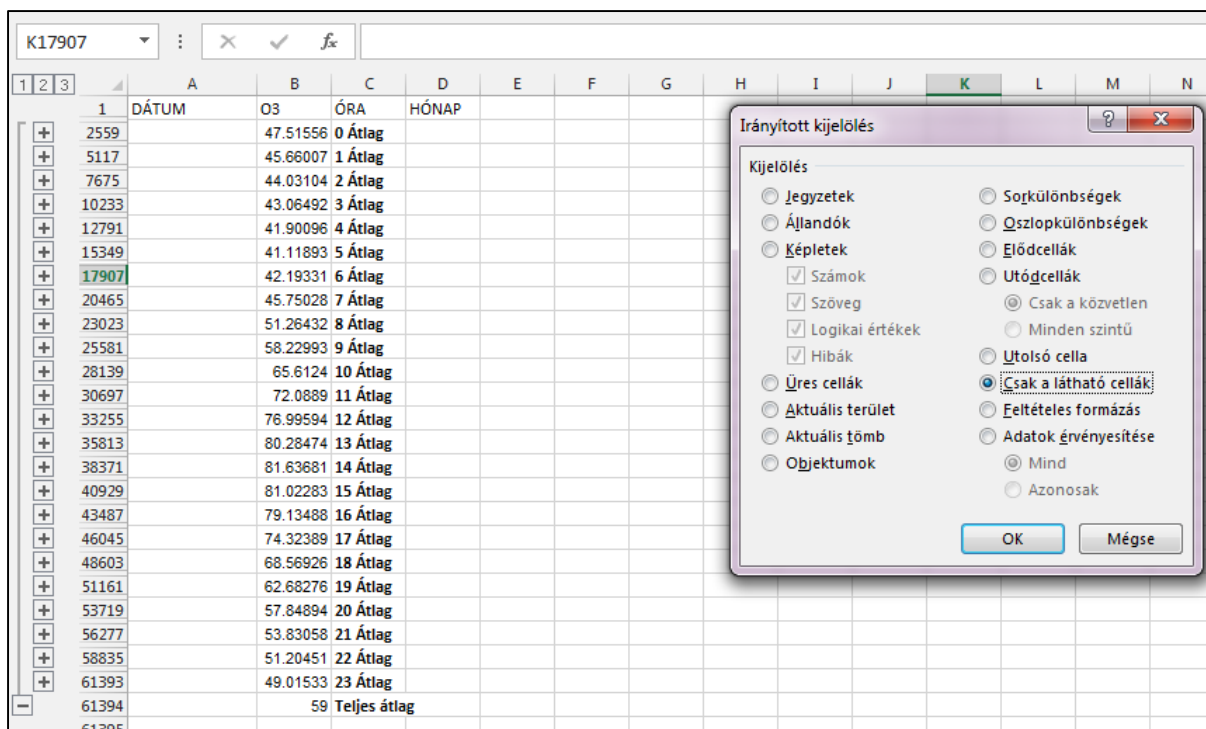
20. ábra: Az óránkénti átlagok és a teljes átlag [2]

A [3]-as esetén pedig minden adatot és az azonos értékkel rendelkező csoportosító változók után beszűrve az átlagokat. A 00:00-órás átlagokat pl. a 2559-es sorban láthatjuk, hiszen az összes adatból (61368 db) a 2559-es sorba esik a 00:00-nak megfelelő 7 év * 365 nap + 2 szökőnap + 1 sor fejléc, és ezek alá kerül beszűrással a részösszeg.

Mivel nekünk az átlagok kellenek, a [2]-es szintet fogjuk használni.

Irányítottan kijelöljük a továbbiakban használni kívánt adatokat: használjuk a „KEZDŐLAP” fülön „Keresés és kijelölés” gombot, az „Irányított kijelölés” panelen „Csak a látható cellák” kiválasztásával szűrjük ki, majd CTRL+C-vel másoljuk ki a látható cellákat (21. ábra).

Erre azért van szükség, mert noha nem látjuk az összes mért adatot a táblázatban, mint a [3]-as szint esetén, de attól még azok ott vannak, és nincs rájuk szükségünk a havi átlagok menetének képzésénél a továbbiakban, csak a [2]-es szint látható celláiból kell dolgoznunk.



21. ábra: A látható cellák kijelölése

A kimásolt cellákat irányítottan beillesztjük egy új munkalapra. Az új munkalap A1-es cellájában „Jobb kattintás/Irányított beillesztés/Értéket kiválasztása” műveletsort hajtjuk végre.

Azért kell értéként beilleszteni, mert maga részösszeg is egy függvény, és a továbbiakban nincs szükség a függvényhivatkozásra, csak magára az óránkénti átlagokra, azaz csupán a függvényértékekre. Ezzel a táblázatot számunkra megfelelő formátumúra alakítottuk. Ehhez kitöröljük a 26-os sort, mert nem lesz szükségünk a teljes átlagra. Átírjuk a „DÁTUM” fejlécet az A1 cellában „ÓRÁ”-ra, tekintve, hogy óránkénti átlaggal fogunk dolgozni (22. ábra).

Kivesszük az átlagok sorszámát a C oszlopból, ahol az szöveggént van jelen (balra zárt tartalom):

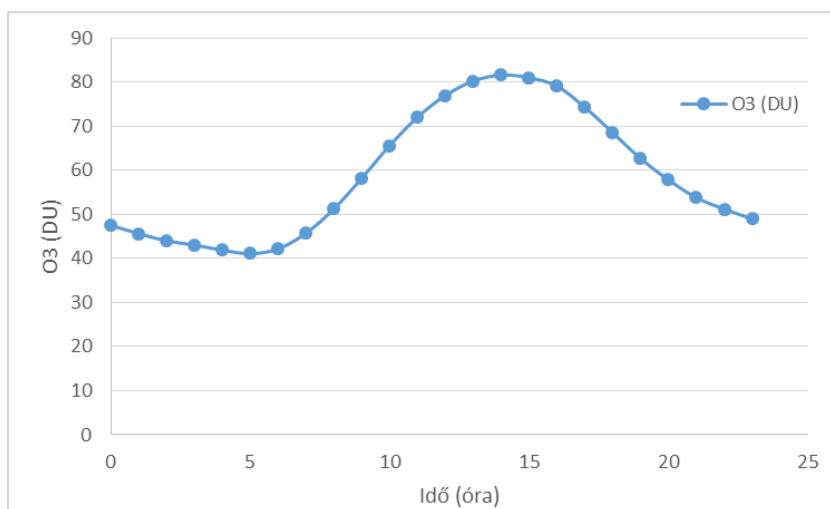
A2→ =ÉRTÉK(BAL;C2;2)

Ezt lemásoljuk a táblázat aljáig, vagyis az A25 celláig. A most létrehozott A2:A25 cellákat kijelöljük és irányítva beillesztjük „önmagukra”, mivel következő lépésben a C oszlopot ki fogjuk törölni, így az A oszlopban lévő függvényhivatkozásunk nem lenne a továbbiakban értelmezhető és „elveszítenénk” az óra-értékeket.

A1				
	A	B	C	D
1	DÁTUM	O ₃ (DU)	ÓRA	HÓNAP
2		47.51556	0 Átlag	
3		45.66007	1 Átlag	
4		44.03104	2 Átlag	
5		43.06492	3 Átlag	
6		41.90096	4 Átlag	
7		41.11893	5 Átlag	
8		42.19331	6 Átlag	
9		45.75028	7 Átlag	
10		51.26432	8 Átlag	
11		58.22993	9 Átlag	
12		65.6124	10 Átlag	
13		72.0889	11 Átlag	
14		76.99594	12 Átlag	
15		80.28474	13 Átlag	
16		81.63681	14 Átlag	
17		81.02283	15 Átlag	
18		79.13488	16 Átlag	
19		74.32389	17 Átlag	
20		68.56926	18 Átlag	
21		62.68276	19 Átlag	
22		57.84894	20 Átlag	
23		53.83058	21 Átlag	
24		51.20451	22 Átlag	
25		49.01533	23 Átlag	
26				

22. ábra: Az óránkénti átlagok beillesztése után látható kép

Most kitöröljük a C és D oszlopokat, mivel azokra a grafikon létrehozásához már nincs szükség. Végző lépésként kijelöljük az A1:B25 cellatartományt és beszúrunk egy pont x-y grafikont. A lépések: „Beszúrás fül/Diagramok/Pont/Pont vonalakkal és jelölőkkel” típus. A diagramot megfelelően formázzuk, töröljük a jelmagyarázatot, feliratozzuk a tengelyeket: a mértékegységet jelölje az X tengely esetén óra, az Y tengely esetén O₃ (DU) (23. ábra).



23. ábra: Az O₃ koncentráció napi menete

Értelmezzük a kapott eredményeket. A grafikonról leolvasható, hogy a troposzférikus ózonkoncentráció délután 14:00-kor éri el a maximumát (81,63 DU), ami nem meglepő, hiszen eddigre érte el a nitrogén-dioxid és a napsugárzás is a maximumát. Az ezt követő csökkenés pedig a napsugárzás intenzitás csökkenésének tudható be. Ismeretes, hogy a délutáni csúcsforgalom egy ismételt nitrogén-dioxid csúcsot eredményez, de ehhez már ózoncsúcs nem tartozik a csökkent napsugárzás miatt.

(b) feladat:

Az óránkénti átlagos menet képzése után az előbbiekhöz hasonlóan hozzuk létre a havonkénti menetet, hogy meglássuk, hogyan változnak a hónapok átlagai egy éven belül. Távolítsuk majd el az eredeti adattáblából a részösszegeket és képezzük a havonkénti átlagokat.

Kijelöljük az eredeti adathalmazt az „Munka 1” munkalapon (CTRL+A), majd kiválasztjuk az „Adatok/Részösszeg” parancsát. Itt el kell távolítanunk az összes részösszeget az „Összes eltávolít” gombra kattintva, és utána újraprendezni az adatsort HÓNAP szerint, hiszen havonkénti átlagos menetet szeretnénk látni a következőkben.

Rendezzük sorba először a mostani részösszeg-mentesítéssel visszkapott adatokat HÓNAP szerint, majd képezzük a havonkénti átlagokat részösszeggel; utóbbit az alábbiak szerint:

Csoportosítási alap: HÓNAP

Melyik Függvénnyel: Átlag

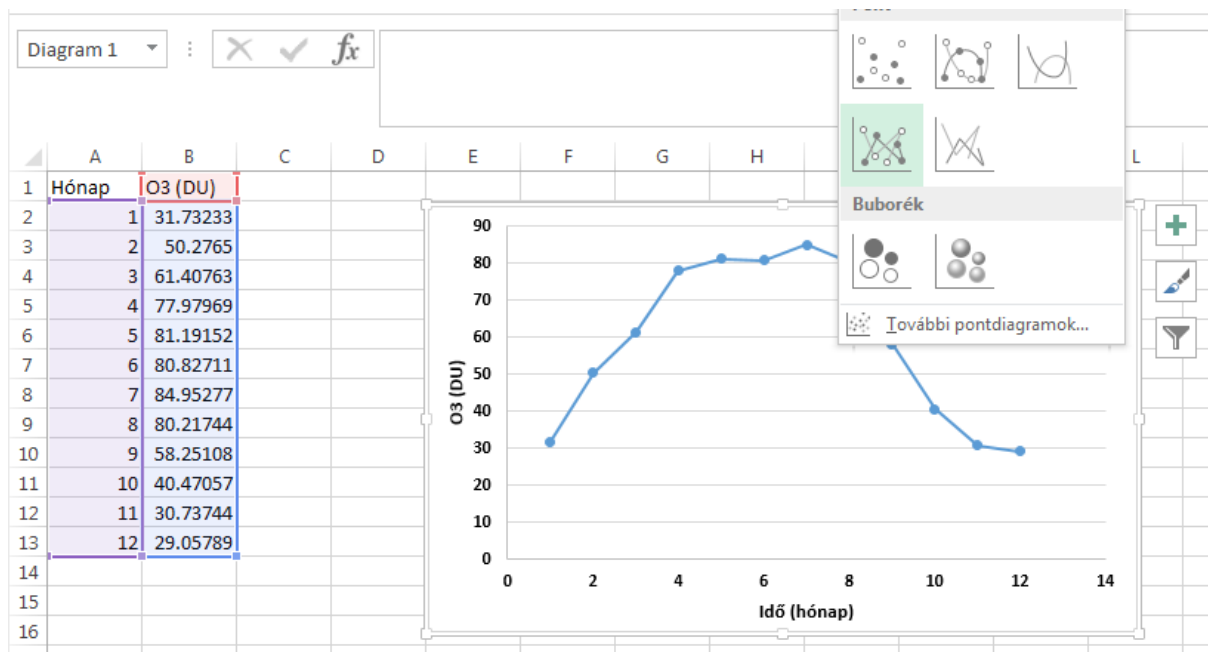
Összegzendő oszlopok: O3 (DU)

A kapott táblázatot [2]-es szinten jelenítjük meg, és az eredményt a korábbiakhoz hasonlóképpen, irányítottan kimásoljuk, és irányítottan értékként beillesztjük egy új munkalapra, mint a korábbiakban.

Ugyancsak a korábbiak mintájára átalakítjuk a számunkra megfelelő formátumúra a következőképpen:

A2 → =ÉRTÉK(BAL(D2;2))

majd továbbmásoljuk az A3:A13 cellákba. Ezután másoljuk az A2:A13 cellákat vágólappra, majd irányított beillesztéssel ismét illesszük a cellák tartalmát saját magukra. Végül töröljük ki a nem szükséges információt és ábrázoljuk a kapott értékeket, megfelelően formázva a diagramot (24. ábra).



24. ábra: Az O₃ koncentrációjának éves menete (havi átlagokból)

Értelmezzük a kapott eredményeket. Az elméleti háttér ismeretében nem meglepő, hogy a nyári hónapokban éri el a troposzférikus ózon a maximumát, mivel ekkor a legnagyobb a napos órák száma, így a fotokémiai aktivitás is. Ez látszik a fenti diagramról is. Áprilistól augusztusig veszi fel a legmagasabb és az ahhoz közeli értékeket.

Összefoglalva: a feladat bemutatta, hogyan lehet egy idősből hatékonyan átlagokat (vagy bármilyen más részösszeget) képezni; gyakorlatban pedig láthattuk, hogy amit elméletben tanultunk a fotokémiai, Los Angeles típusú szmogról, az például egy Fertő tóhoz közeli meteorológiai állomás troposzférikus ózon adatain is tapasztalható.

8. Keresztkorreláció és autokorreláció

a. Keresztkorreláció

Példafájl neve: Rajka.xls

A feladat célja:

Bemutatni a keresztkorreláció számításának menetét, majd annak segítségével meghatározni, hogy a Duna vízszintváltozása mennyi idő múlva érezteti hatását egy megfigyelőkút vízszintjében. Ez nem azt jelenti, hogy egy vízmolekula ennyi idő alatt ér el a Dunától a figyelt pontig, hanem azt, hogy az árvíz nyomáshulláma mennyi idő alatt éri el a kutat.

Új ismeretek:

- Korrelációs együttható számítása
- KORREL függvény használata
- Keresztkorreláció függvény létrehozása

Ebben az alfejezetben a keresztkorrelációval foglalkozunk, a következőben pedig az alkalmazott autokorrelációval, amely szorosan összefügg az itt bevezetett fogalmakkal. Ezért célszerű ebben a fejezetben együtt tárgyalni őket.

Elméleti háttér

- Korrelációs együttható:

A korrelációs együttható két változó közötti lineáris kapcsolatot mér, dimenzió nélküli szám. A becslt korrelációs együttható jele 'r'; az x_i és y_i egy vizsgálat során mért két adatot, adatpárt jelent. n darab megfigyelés esetén

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

ahol: $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ és $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$.

A korrelációs együttható -1 és +1 közötti értéket vehet fel. Minél közelebb van az érték -1-hez vagy +1-hez, annál inkább lineárishoz közelebb a kapcsolat a két változó között. Tökéletes lineáris kapcsolat esetén, azaz amikor minden pont egy egyenesen van akkor pontosan a +1 vagy -1 értéket veszi fel. Amennyiben a korrelációs együttható értéke nulla (vagy ahhoz közeli szám), az nem azt jelenti, hogy nincs kapcsolat a két változó között, mindösszesen azt jelenti, hogy nincs közöttük, vagy csak nagyon gyenge a lineáris kapcsolat. Két független változó esetén a korrelációs együttható értéke nulla. Fordítva ez nem igaz, mert a lineáris kapcsolat hiánya még nem zárja ki a négyzetes, exponenciális, stb. kapcsolat lehetőségét. Nulla korrelációs együttható esetén a változók tehát korrelálatlanok, de nem függetlenek.

- Autokorreláció:

Az autokorreláció egyazon változó önmagához képest időben eltolt értékei közötti lineáris kapcsolatot vizsgálja. Az $(x_{k+1}, x_1); (x_{k+2}, x_2); \dots; (x_n, x_{n-k})$ párok korrelációját számolva kapjuk a k egységhez tartozó autokorreláció becült értékét. Ezen eljárás stacionárius megfigyelésekre alkalmazható.

- Keresztkorreláció:

A keresztkorreláció két különböző változó lineáris kapcsolatának időbeli változását vizsgálja. Az egyik adatsor értékeit összevetjük a másik adatsor k egységgel eltolt értékeivel, majd ezek korrelációját számoljuk. Azaz a $(x_{k+1}, y_1); (x_{k+2}, y_2); \dots; (x_n, y_{n-k})$ párok korrelációját számolva kapjuk a k egységhez tartozó keresztkorreláció becült értékét. Hasonlóképpen lehet vizsgálni a keresztkorrelációt akkor, ha az y adatsor van időben előrébb. Összességében a keresztkorrelációk számításával az adatsorban a késleltetett hatásokat figyelhetjük meg.

Feladatban felhasznált függvény:

- **KORREL függvény:** Két adathalmaz korrelációs együtthatóját számolja ki az alábbi formában. A függvény a tömb1 és tömb2 cellatartományok korrelációs együtthatóját adja eredményül.

$$= \text{KORREL}(\text{tömb1}; \text{tömb2})$$

Feladat:

Az első – „kút127_2002” nevű – munkalap A oszlopában láthatjuk azokat az időpontokat, amikor a Duna, valamint a Dunától néhány km távolságban levő megfigyelőkút vízszintadatait leolvasták. Ilyen vizsgálatoknál az adatoknak dátum szerint kell sorba rendezetteknek lenniük, tehát ezt mindig ellenőriznünk kell, ahogy azt is, hogy a mintavétel ekvidisztáns-e (azonos időközű-e).

A B és C oszlopokban a megfigyelőkút, valamint a Duna vízszintadatai jelennek meg. Van, ahol üres cellával találkozunk, Ez azt jelenti, hogy abban az időpontban azon a helyen nem végeztek vízszintmérést. A vizsgálat szempontjából azonban nem okoz gondot, ha adathiány fordul elő ezekben az oszlopokban, mivel az érintett „adattípust” az Excel kihagyja a korreláció számításából.

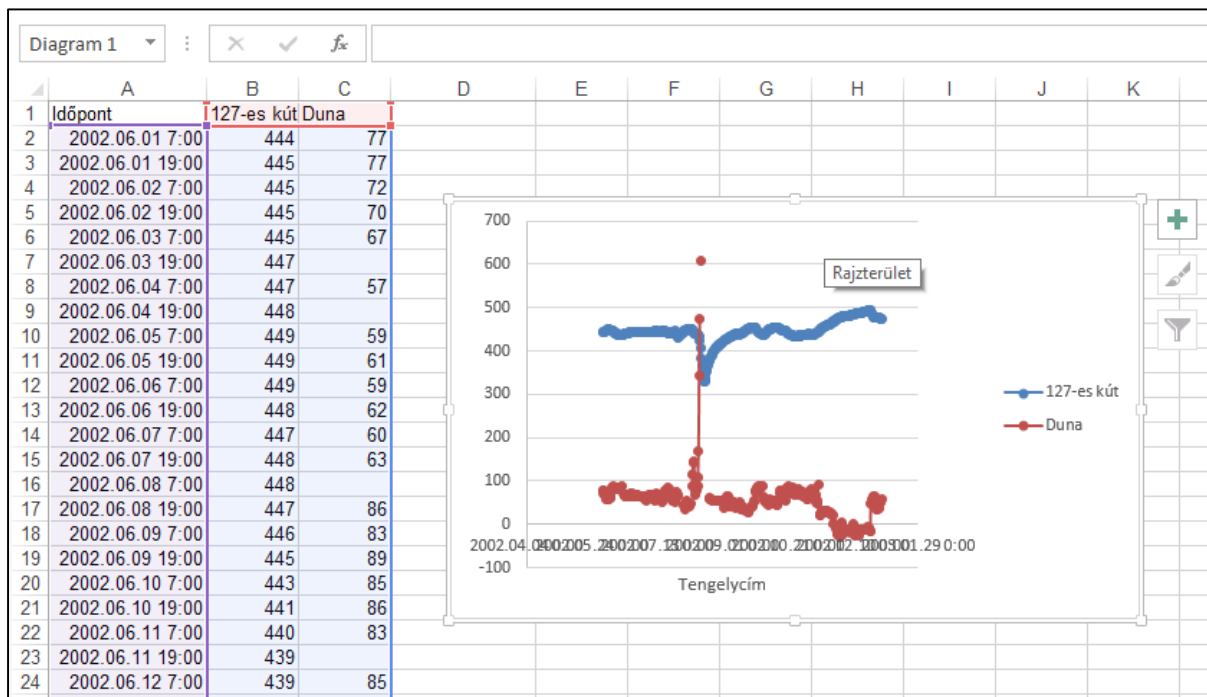
- (a) A feladatban arra keressük a választ, hogy a Dunán érkező árvíz nyomásfrontja mennyi idő alatt ér el a 127-es megfigyelőkúthoz, más szavakkal, a folyó vízszintváltozása mennyi idő elteltével változtatja meg a megfigyelőkút vízszintjét.

Megoldás:

(a) feladat:

Első lépésként az adatsort magát érdemes vizuálisan megvizsgálunk, hogy lássuk a vízszintekben bekövetkezett változásokat. Az adatsor 12 óránként mutatja a vízszintértékeket 2002.06.01. 7 órától 2002.12.31. 19 óráig, így kapjuk a 428 adatot (214 nap*naponta 2 mérés). Ábrázoljuk grafikonon a teljes adatsort: jelöljük ki az adatokat (A1:C429), majd a

„Beszúrás” menüpontban a „Pont xy grafikon”-ok közül válasszuk ki a felső sorban a középső grafikon
(25. ábra).



25. ábra: A 127-es kút és a Duna vízszintidősora

A grafikonról, és az adatokból együttesen leolvasható, hogy augusztus elején a Duna vízszintje erősen megemelkedett, Rajkánál meghaladta a 6 métert, majd újra lecsökkent. Ezzel szemben a 127-es kútban a vízszint ellentétesen viselkedett. Mi lehetett ennek a jelenségnek az oka?

Míg a Duna vízszintjét a mederben elhelyezett vízszintmérőléccel segítségével mérték (ez egy konstans értékkel tér el a tengerszint feletti magasságtól), addig a kút adatai a csőperemtől mért távolságból képződtek. Így fordulhat elő az ellentétes menet, míg a gyakorlatban természetesen a Duna vízszintemelkedése a kútban is vízszintemelkedést okoz. Ebből a példából jól látszik, hogy adatelemzéskor fontos, hogy tisztában legyünk az adatok háttérével.

Összegezve: kitűnik, hogy az árhullám a Dunán 2002.08.15-én érkezett meg a rajkai állomáshoz, és 2002.08.18-án a 127-es kúthoz. Az árhullám levonulása ezután következett be. Mivel két külön jelenségről beszélünk – hiszen más-más hidrogeológiai folyamattal írható le a közetváz feltöltődése egy árhullám kialakulása során, és leürülése egy árhullám levonulása után – ezért szét kell választanunk a két időszakot.

Jelöljük ki az adattáblát a fejléctől kezdve a 2002.08.18-as nap esti méréséig (A1:C159) és másoljuk át egy új munkalapra, majd nevezzük el a munkalapot „1.fele” néven.

A második – az „1.fele” nevű – munkalapon levő adatsoron végezzük el a továbbiakban a keresztkorreláció-vizsgálatot. Az adatsor itt már csak 2002.08.18. 19 óráig tart, 158 adatot tartalmaz. Mivel a Dunának a kútra gyakorolt hatását vizsgáljuk, ezért fontos, hogy a B oszlopban a 127-es kút adatai legyenek, a C-ben pedig a Duna vízszintjei. Nem mindegy, hogy melyik adatsort melyikhez képest toljuk el időben. A mi feladatunk, hogy azt modellezzük, melyik időeltolásnál éri el maximumát a korrelációs együttható.

Első lépésként jelöljük ki a Duna adatait (C2:C159) és „toljuk el” az adatokat 12 órával a kútadatokhoz képest! (A kijelölésnél ügyelni kell az üres cellákra amennyiben a CTRL+NYÍL kombinációt használjuk. Az ezzel kapcsolatos részletekről a következő alfejezetben olvashatunk). Ez azt jelenti, hogy 12 óránként tudjuk megvizsgálni a korrelációt az adatsorok között. CTRL+C billentyűkombinációval tegyük ki a vágólapra, majd a CTRL+V paranccsal másoljuk be a kijelölt C2:C159 cellákat a D3,E4,F5,G6,H7,I8,J9,K10,L11 cellákkal kezdődően és azok alá. Így összesen $9 \times 12 = 108$ órával töltük el a Duna adatsorát az eredetihez képest.

Ha megnézzük az idősor végét, láthatjuk, hogy az eltolás következtében a 160. sortól nagyobb sorszámú sorokba is kerültek adatok, amelyekre azonban nincs szükségünk, ezeket ki kell törölni:

Bal egérgombbal jelöljük ki a 160. sort. CTRL+SHIFT+END billentyűkombinációval jelöljük ki a kitörlendő cellákat, majd DELETE gombbal töröljük (26. ábra).

A160														
	A	B	C	D	E	F	G	H	I	J	K	L		
144	2002.08.11 7:00	439	77	67	146	142	114	87		48		42		
145	2002.08.11 19:00	439	81	77	67	146	142	114	87		48			
146	2002.08.12 7:00	438	78	81	77	67	146	142	114	87		48		
147	2002.08.12 19:00	438	87	78	81	77	67	146	142	114	87			
148	2002.08.13 7:00	437	108	87	78	81	77	67	146	142	114	87		
149	2002.08.13 19:00	436	170	108	87	78	81	77	67	146	142	114		
150	2002.08.14 7:00	433	344	170	108	87	78	81	77	67	146	142		
151	2002.08.14 19:00	425	475	344	170	108	87	78	81	77	67	146		
152	2002.08.15 7:00	408		475	344	170	108	87	78	81	77	67		
153	2002.08.15 19:00	385	609		475	344	170	108	87	78	81	77		
154	2002.08.16 7:00	367		609		475	344	170	108	87	78	81		
155	2002.08.16 19:00	350			609		475	344	170	108	87	78		
156	2002.08.17 7:00	335				609		475	344	170	108	87		
157	2002.08.17 19:00	331					609		475	344	170	108		
158	2002.08.18 7:00	331						609		475	344	170		
159	2002.08.18 19:00	338							609		475	344		
160										609		475		
161											609			
162												609		

26. ábra: Az adatsor eltolása után a felesleges adatok törlése

Következő lépésként vizsgáljuk meg a 127-es megfigyelőkút, valamint a Duna „eltolt” adatsorai közötti korreláció mértékét „eltolásonként” (estünkben 12 óránként). Ezt úgy tudjuk megtenni, ha kiszámoljuk a kút adatai és mindegyik eltolás közötti korrelációs együtthatók nagyságát. Az együtthatók értékeiből következtethetünk majd a Duna hatásának mértékére. Ennek menetét a következő lépések adják:

B163 → = „eltolás mértéke (óra)”,

B164 → = „korrelációs eh.”,

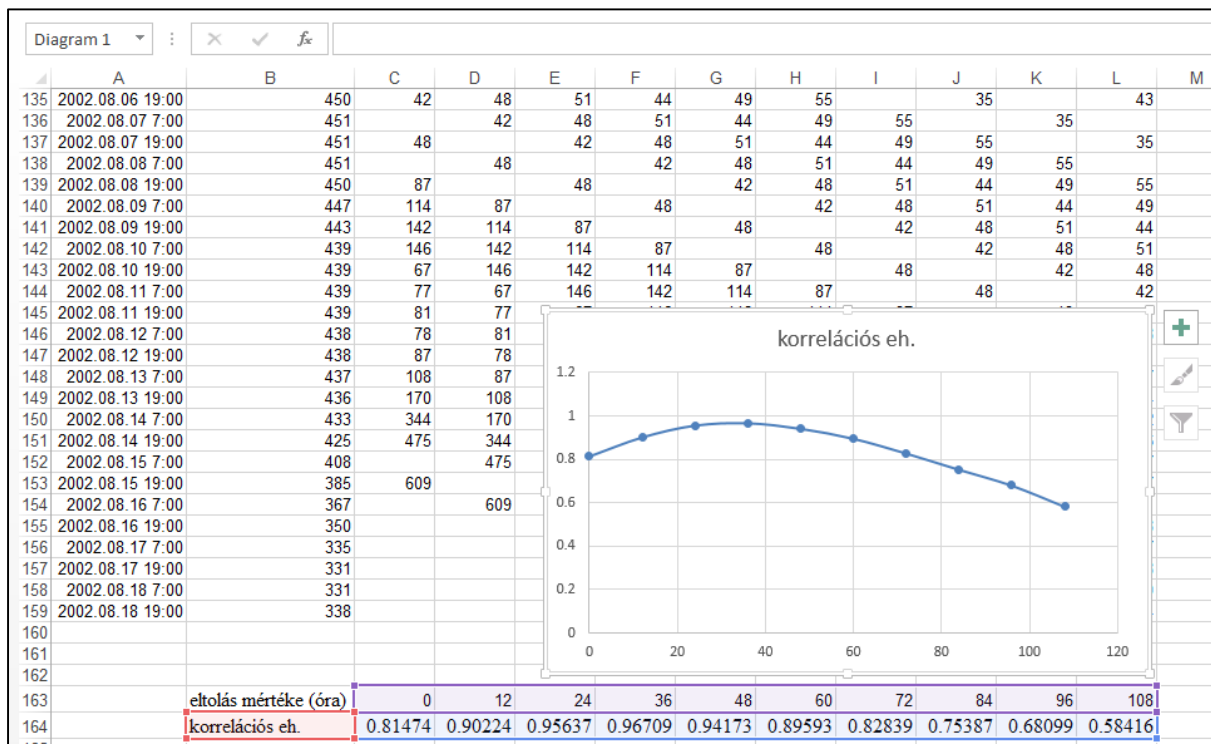
C163:L163 → = 0,12,24,...,108. 12 órás eltolások 0 órától 108 óráig.

C164 → =KORREL(\$B2:\$B159;C2:C159)*-1

Ezt másoljuk tovább az L164 celláig. A -1-es szorzás azért célszerű, mert ahogy az előzőekben láttuk, az eredeti adatsorban a kút vízszintje ellentétesen mozog a Duna

vízszintjéhez képest. Mivel csak a Duna adatait töltük el, ezért az „eltolt” Duna-adatokat mindig a kút vízszintjét adó B oszloppal kell korreláltatnunk. Ez magyarázza azt, hogy a függvényben a B oszlopot rögzíteni kellett a \$ jellel.

Ábrázoljuk a kapott korrelációs együtthatókat az időeltolás mértékének függvényében: Ehhez jelöljük ki a B163:L164 cellákat, majd a „Beszúrás fül/Diagramok”/Pont (X,Y)” menüpontból, a vonallal összekötött grafikonok közül a felső középsőt válasszuk ki (27. ábra).



27. ábra: A Duna vízszintje és a talajvízszint keresztkorrelációja

Értelmezzük a kapott eredményeket: a grafikonról leolvasható, hogy a Duna vízszintjében bekövetkezett emelkedés kb. 36 óra elteltével érezteti a legnagyobb mértékben a hatását a tőle távolabb lévő talajvízszint-megfigyelőkút vízszintjére. Ezt követően a Duna hatása csökken. Miután 12 óránként végezték a méréseket, a Duna hatását is csak „valahányszor” 12 órában tudjuk megadni. Gyakoribb vízszintmérés esetén pontosabb eredményt kapnánk.

b. Autokorreláció

Példafájl neve: ozon.xls

Adatforrás: <http://www.nilu.no/projects/ccc/onlinedata/ozone/>. Ez a honlap az EMEP mérőhálózat interneten elérhető adatait tartalmazza, több tíz európai mérőállomásról különböző hosszúságú időintervallumokra.

A feladat célja:

Megvizsgálni, hogy az óránként mért ózonnak milyen az „emlékezete”. Bemutatni az autokorreláció-vizsgálat menetét egy ózon-adatsoron, miközben makró használatával „automatikusan”, több adatsorra hajtjuk végre a folyamatot.

Új ismeretek:

- Autokorreláció függvény létrehozása
- Makrókészítés

Elméleti háttér:

Ld. 7. fejezetet.

Ld. Keresztkorreláció alfejezetet.

Feladat:

- (a) Határozzuk meg az egyes mérési pontokon az ózonkoncentráció idősorának autokorreláció függvényeit és hasonlítsuk össze a kapott eredményeket!

Megoldás:

(a) feladat:

Töltsük be az **ozon.xls** fájlt. A fájl nyolc munkalapot tartalmaz, mindegyiken különböző országokban mért troposzférikus ózonkoncentráció-adatsorokat láthatunk. Az első munkalapon található adathalmazzal már találkoztunk a 7. fejezetben.

A feladat célja, hogy a mérési pontokon meghatározzuk az ózonkoncentráció idősorának autokorreláció függvényeit és összehasonlítsuk a kapott eredményeket. Az Excelben lehetőség van makró írására és rögzítésére Visual Basic programozási nyelven (LIBERTY, 2003), melynek segítségével egy munkafolyamatot elegendő csak egyszer lépésenként végigvinni. Ezt követően a makróban „rögzített” műveleteket a többi adatsoron egyszerűen csak le kell futtatni.

Fontos megjegyzés: ahhoz, hogy egy makrót több adatsoron alkalmazhassunk, az adatsoroknak hasonló szerkezetűnek kell lenniük. Nézzük meg az első munkalap („Illmitz, A”) felépítését:

Az A oszlopban a dátum szerepel, a B oszlopban pedig az ózonkoncentráció-értékek, dátum szerint sorba rendezve. 1990.01.01-től 1996 végéig történtek óránkénti mérések, így 61368 adat tartozik egy idősorba ((7 év * 365 nap +2 nap) * naponta 24 mérés). Az oszlopok fejlécet nem tartalmaznak.

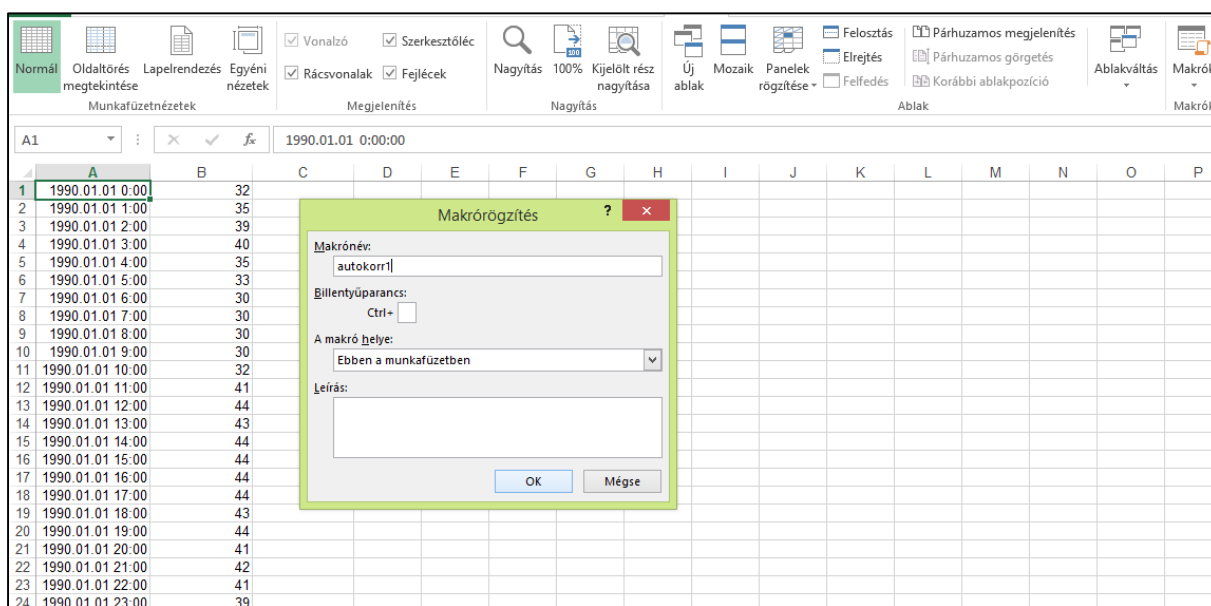
Ezt a felépítést követi a többi munkalap idősora is. A keresztkorreláció vizsgálatához is – hasonlóan az autokorrelációhoz – ekvidisztáns (azonos időközű) mintavételre van szükség, valamint a koncentráció-értékekben az adathiány itt is megengedett (természetesen azzal a megszorítással, hogy a számított autokorrelációk kevesebb mintapárból készülnek).

Az autokorreláció vizsgálatának menete nagyrészt megegyezik a keresztkorreláció lépéseivel, miszerint: (i) a mintavételi gyakorisággal (ebben az esetben ez 1 óra) eltoljuk a troposzférikus ózonkoncentráció adatsorát önmagához képest, majd (ii) kiszámítjuk az eredeti és az „eltolt” adatsorok közti korreláció mértékét. A korrelációs együtthatók segítségével előállítjuk az autokorrelációs együtthatók idősorát. Grafikonon szemléletesen ábrázolhatjuk, hogy a folyamat adott idő elteltével mennyire autokorrelált. Mind a három mérési helyszín adatsorából meg fogjuk határozni az ózonkoncentrációk autokorreláció függvényeit.

Ahhoz hogy a műveletsorozatot megsokszorozzuk, célszerű az első munkalapon („Illmitz, A”) elvégezni és rögzíteni a munkamenetet, ezért válasszuk ki az első munkalapot.

Első lépésként el kell indítani a makró-„felvételt” a következőképpen: „Nézet lapfűl/Makrók csoport/Makró rögzítése”. A feladatot a könnyebb áttekinthetőség kedvéért két makró segítségével fogjuk megoldani. Az elsővel eltoljuk az adatsort, a másodikkal pedig kiszámoljuk az eltolásokhoz tartozó korrelációs együtthatók értékeit.

Adjunk nevet az első makrónak: a makrónév legyen „autokorr1”, majd az OK gombbal indítsuk el a makró felvételét **(28. ábra)**. Lehetőségünk van billentyűkombinációt hozzárendelni a létrehozott makróhoz, illetve egy leírást is készíthetünk arról, hogy miről szól az adott makró. Ezekről itt most eltekintünk.



28. ábra: Makrófelvétel indítása

Fontos megjegyezni, hogy innentől kezdve a makró rögzíti minden lépésünket, tehát célszerű a szükséges feladatokat előre megtervezni és pontosan végrehajtani, mert amennyiben téves műveletet végzünk, előlről kell kezdeni a makró felvételét. Itt meg kell jegyeznünk, hogy az Excel a makrókat magát Visual Basic (VBA) programozási nyelven rögzíti, tehát aki rendelkezik ilyen jellegű programozási ismeretekkel, az a felvétel végén tud „módosítani a műveleteken” (ajánlott irodalom: KOVALCSIK, 2010).

CTRL+HOME billentyűkombinációval ugorjunk az A1 cellába. Az oszlopok nem tartalmaznak fejléceket, ezért szúrjunk be egy sort az első sor elé (bal egérgombbal jelöljük ki az első sort, majd jobb egérgomb után „Beszúrás” parancs) és nevezzük el az oszlopokat:

A1 → = „dátum”,

B1 → = „O3 (DU)”.

A következő lépés, hogy a B oszlopban levő koncentráció-értékeket eltoljuk az eredeti adatsorhoz képest. Lényeges különbség a keresztkorrelációhoz képest, hogy itt egy „0” órás eltolást is létrehozunk, és csak ezután toljuk el óránként az adatokat.

Jelöljük ki az ózonkoncentráció értékeit tartalmazó B2:B61369 cellákat. Ezt gyorsan a következőképpen tehetjük meg: Az A oszlop bármelyik cellájából kiindulva END+↓ gombokkal a 61369. sorba jutunk. Azért fontos az A oszlop, mert abban nincs adathiány, így valóban az adatsor végére ugrunk. A B61369. cellából kiindulva a CTRL+SHIFT+HOME paranccsal jelöljük ki a teljes A1:B6139 adatsort. A SHIFT gombot nyomva tartva a →, majd ↓ nyíllal érhetjük el a kívánt B1:B6139 tartomány kijelölését. (Megjegyezzük, ez a kijelölés más módon, például egér segítségével is elvégezhető.)

A kijelölt B1:B6139 tartományt a CTRL+C billentyűkombinációval tegyük ki a vágólapra, majd a CTRL+V billentyűkkel másoljuk be a C2,D3, E4, F5,...,AG32 cellákba. Így összesen 30 órával toltuk el az adatsort az eredetihez képest (29. ábra).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	dátum	O3 (DU)													
2	1990.01.01 0:00	32	32												
3	1990.01.01 1:00	35	35	32											
4	1990.01.01 2:00	39	39	35	32										
5	1990.01.01 3:00	40	40	39	35	32									
6	1990.01.01 4:00	35	35	40	39	35	32								
7	1990.01.01 5:00	33	33	35	40	39	35	32							
8	1990.01.01 6:00	30	30	33	35	40	39	35	32						
9	1990.01.01 7:00	30	30	30	33	35	40	39	35	32					
10	1990.01.01 8:00	30	30	30	30	33	35	40	39	35	32				
11	1990.01.01 9:00	30	30	30	30	30	33	35	40	39	35	32			
12	1990.01.01 10:00	32	32	30	30	30	30	33	35	40	39	35	32		
13	1990.01.01 11:00	41	41	32	30	30	30	30	33	35	40	39	35	32	
14	1990.01.01 12:00	44	44	41	32	30	30	30	30	33	35	40	39	35	32
15	1990.01.01 13:00	43	43	44	41	32	30	30	30	30	33	35	40	39	35
16	1990.01.01 14:00	44	44	43	44	41	32	30	30	30	30	33	35	40	39
17	1990.01.01 15:00	44	44	44	43	44	41	32	30	30	30	30	33	35	40
18	1990.01.01 16:00	44	44	44	44	43	44	41	32	30	30	30	30	33	35
19	1990.01.01 17:00	44	44	44	44	44	43	44	41	32	30	30	30	30	33
20	1990.01.01 18:00	43	43	44	44	44	44	43	44	41	32	30	30	30	30
21	1990.01.01 19:00	44	44	43	44	44	44	44	43	44	41	32	30	30	30
22	1990.01.01 20:00	41	41	44	43	44	44	44	44	43	44	41	32	30	30
23	1990.01.01 21:00	42	42	41	44	43	44	44	44	44	43	44	41	32	30
24	1990.01.01 22:00	41	41	42	41	44	43	44	44	44	44	43	44	41	32
25	1990.01.01 23:00	39	39	41	42	41	44	43	44	44	44	44	43	44	41
26	1990.01.02 0:00	39	39	39	41	42	41	44	43	44	44	44	44	43	44

29. ábra: Az ózon-adatok eltolása

Ha megnézzük az adatsor végét, láthatjuk, hogy az eltolás következtében a 61370. sortól kezdve is képződtek adataink, amikre azonban nincs szükségünk, így ezeket kitöröljük: bal egérgombbal jelöljük ki a 61370. sort. CTRL+SHIFT+END billentyűkombinációval jelöljük a kitörölendő cellákat, majd DELETE gombbal töröljük.

Amikor idáig eljutottunk, állítsuk le a makrót: „Nézet fül/Makrók gomb/Rögzítés vége”.

Ekkor megtehetjük, hogy az eddig elvégzett és a makróban rögzített lépéseket végrehajtjuk a többi adatsoron. Nyissuk meg a második munkalapot („Schauinsland, DE”), majd a „Nézet

fül/Makrók gomb/Makrók megjelenítése” ablakban válasszuk ki a futtatni kívánt makrót, és az Indítás paranccsal futtassuk. Ezt követően a többi munkalapon is hasonlóan járunk el.

Következő lépésként vizsgáljuk meg az eredeti és az „eltolt” adatsorok közti korreláció mértékét. Ezt úgy tudjuk megtenni, ha kiszámoljuk az eredeti adatok és mindegyik eltolt adatsor közötti korrelációs együtthatókat. Ennek menete a következő:

Lépünk vissza az első munkalapra („Illmitz, A”), és készítsünk egy második makrót:

„Nézet menüpont/Makrók fül/Makró rögzítése”. A makrónév legyen autokorr2, majd OK parancs. Az A oszlop bármelyik cellájából kiindulva End+↓ gombokkal lépünk az adatsor aljára.

B61372 → = „eltolás mértéke (óra)”,

B61373 → = „korrelációs eh.”,

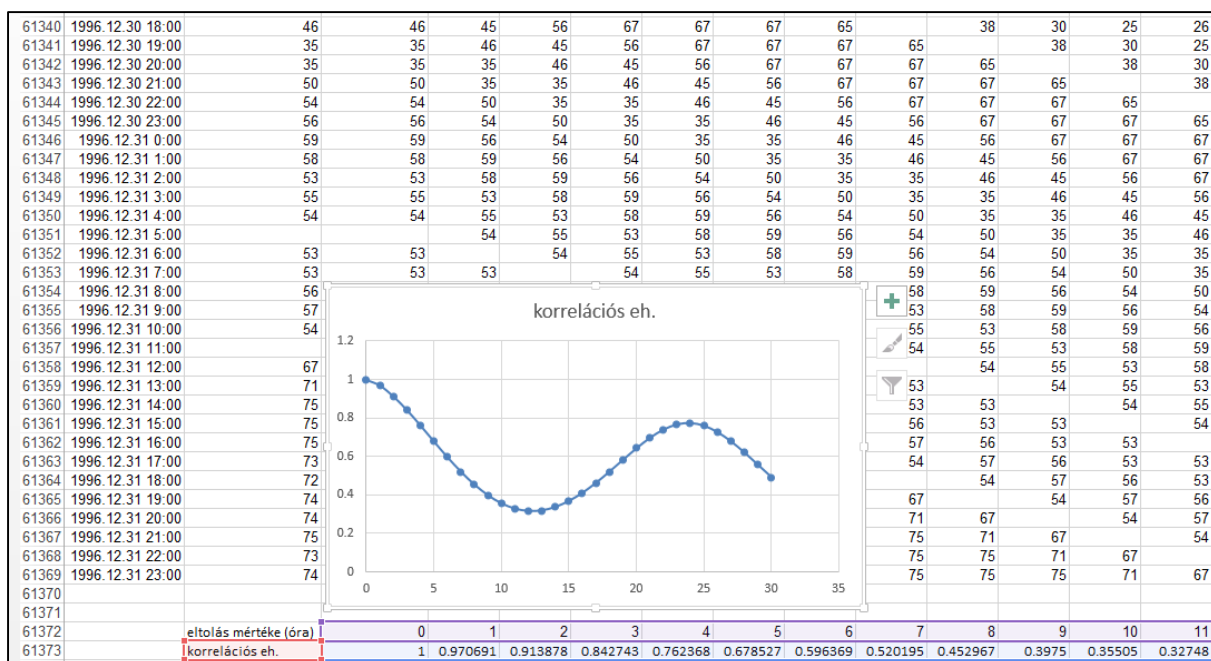
C61372:AG61372 → = 0,1,2,...,30. egy órás eltolások 0 órától 30 óráig.

C61373 → =KORREL(\$B2:\$B61369;C2:C61369).

A C61373 cellát másoljuk tovább a AG61373 celláig. Minden eltoláshoz tartozó adatsort az eredeti, B oszlopban levő adatsorhoz viszonyítjuk, ezért kellett a B oszlopot rögzíteni.

Állítsuk le a makrót: a „Nézet fül/Makrók gomb/Rögzítés vége” paranccsal. Futtassuk le az „autokorr2” makrót a többi munkalapon is.

Ábrázoljuk a kapott korrelációs együtthatókat az eltolás mértékének függvényében valamennyi mérési ponton: jelöljük ki a B61372:AG61373 cellákat. Válasszuk ki a „Beszúrás fül/Diagramok/Pont/vonallal és jelölőkkel összekötött” opciót a grafikonok közül (30. ábra).



30. ábra: Az ózonkoncentrációk autokorrelációja, az „Illmitz, A” állomáson

Összefoglalásként értelmezzük a kapott eredményeket. A grafikonról leolvasható, hogy az Illmitz mérőponton az ózon autokorreláció függvénye periodikus lecsengésű. Megfigyelhető, hogy a más mérőpontokon (szemléletes példák: Schauinsland, és Ispra, IT) a függvény menete eltérő, mutatva, hogy az ózonkoncentráció-idősorok helytől függően különbözően viselkednek.

9. Célértékkeresés

A feladat célja:

A célértékkeresés használatával egy előre meghatározott korrelációs együtthatójú adatsorpár előállítás.

Új ismeretek:

- VÉL és a VÉLETLEN.KÖZÖTT függvények használata
- Célértékkeresés használata

Elméleti háttér:

A természettudományok területén gyakran előfordul, hogy adathalmazunk hiányzó és/vagy hibás adatokat tartalmaz. Ebben az esetben felmerül az adatpótlás szükségessége, lehetősége, amelynek egyik eszköze a lineáris regresszió (KOVÁCS ET AL. 2012c, HATVANI ET AL. 2014). Ilyen jellegű adatpótlást akkor alkalmazhatunk hatékonyan, ha a hiányzó adatot tartalmazó változó más változóval szoros lineáris kapcsolatban van (korrelációs együtthatója 1-hez vagy (-1)-hez közeli érték). Ilyenkor egy X véletlen változó adott x értéke alapján, Y értéket az $E(Y|X=x)$ feltételes várható értékkel becsüljük. Ha ezt minden x-re meghatározzuk, úgy egy elsőfajú regressziós függvényt kapunk. Ha X és Y változók együttes eloszlása normális, akkor az elsőfajú regressziós függvény lineáris. A legkisebb négyzetek módszerével azt a függvényt keressük, amely legjobban közelíti Y változó értékeit. A módszer a mért és a hozzájuk tartozó becsült értékek különbségeinek négyzetösszegét minimalizálja. Ha a két változó együttes eloszlása ismeretlen, akkor is becsülhetjük Y-t ezzel a módszerrel, ekkor másodfajú regressziós függvényt kapunk.

A feladatban használt, véletlen számokat generáló függvények

- **VÉL függvény:** 0 és 1 közé eső véletlen számot állít elő
=VÉL()
- **VÉLETLEN.KÖZÖTT függvény:** Két adott szám közötti véletlen egész számot állít elő
=VÉLETLEN.KÖZÖTT(alsó;felső)

alsó: a függvény által visszaadható legkisebb egész szám.

felső: a függvény által visszaadható legnagyobb egész szám.

Feladat:

- (a) Nyisson meg egy üres munkafüzetet, majd mentse el *célértékkeresés.xlsx* néven! Az A oszlopban hozzon létre 200 db 10 és 300 közé eső véletlen számot!
- (b) A B oszlopban hozzon létre egy olyan, az előző adatsorral megegyező méretű adatsort, amelynek az A oszlopban lévő adatsorral vett korrelációs együtthatója 0,7!

Megoldás:**(a) feladat:**

10 és 300 közé eső véletlen számokat a VÉLETLEN.KÖZÖTT függvénnyel hozhatunk létre azáltal, hogy megadjuk a függvény argumentumában a véletlen számok alsó és felső határait:

A1 → =VÉLETLEN.KÖZÖTT(10;300).

Másoljuk le a függvényt az A200-as celláig.

A kapott véletlen számok a munkalap elemeinek módosítása során újragenerálódnak, így érdemes az előállítást követően a kapott értékeket kijelölni, a vágólapra másolni és önmaguk helyére irányítottan, értéként beilleszteni. Tegyük ezt most meg!

(b) feladat:

A következő feladatunk egy olyan adatsor előállítása, amelynek az A oszlopban lévő adatsorral vett korrelációs együtthatója 0,7. Ehhez először állítsunk elő egy olyan adatsort, amelynek az A oszlopban lévő adatsorral vett korrelációs együtthatója 1. Ebben az esetben a megfelelő értékpárokat egymás függvényében ábrázolva a pontok egy egyenesre esnek. Az egyenes egyenlete $y=ax+b$ alakú, ahol a a regressziós egyenes meredeksége, míg b a tengelymetszet.

Az E1-es cellába írjunk be egy meredekség-értéket (legyen ez 2), az E2-esbe pedig egy tengelymetszet-értéket (az pedig legyen 15):

D1 → Meredekség,

E1 → 2,

D2 → Tengelymetszet,

E2 → 15.

Ezt követően a B oszlopban számítsuk ki az A oszlopban lévő x értékek és az egyenes egyenlete alapján a másik adatsor (y) értékeit. A meredekség és a tengelymetszet értékét tartalmazó cellákat rögzítsük:

B1 → =\$E\$1*A1+\$E\$2.

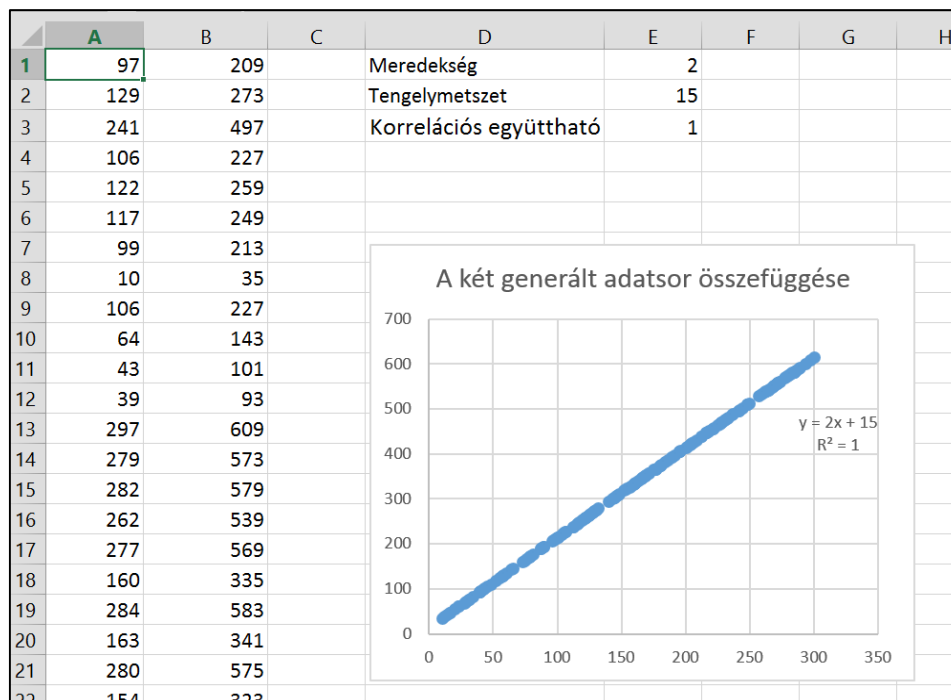
Másoljuk le a B1-es cellában lévő függvényt a B200-as celláig.

Az E3-as cellában számítsuk ki a 2 adatsor korrelációs együtthatóját (az eredmény 1 lesz):

D3 → Korrelációs együttható,

E3 → =KORREL(A1:A200;B1:B200).

Amennyiben egymás függvényében pontdiagramon ábrázoljuk az A és B oszlop értékeit, láthatjuk, hogy a pontjaink a várakozásainknak megfelelően egy egyenesre esnek (**31. ábra**).



31. ábra: A generált adatsor pontdiagramja

A korrelációs együtthatót úgy „ronthatjuk el”, hogy a B oszlop minden értékéhez hozzáadunk egy „ε” véletlen hibát (mindegyik elemhez eltérőt) a VÉL() függvénnyel, mely 0 és 1 közé eső véletlen számokat generál:

C1 → =B1+VÉL().

Másoljuk tovább a C1-es cella tartalmát a C2:C200-as cellákba, majd számítsuk ki az új korrelációs együtthatót:

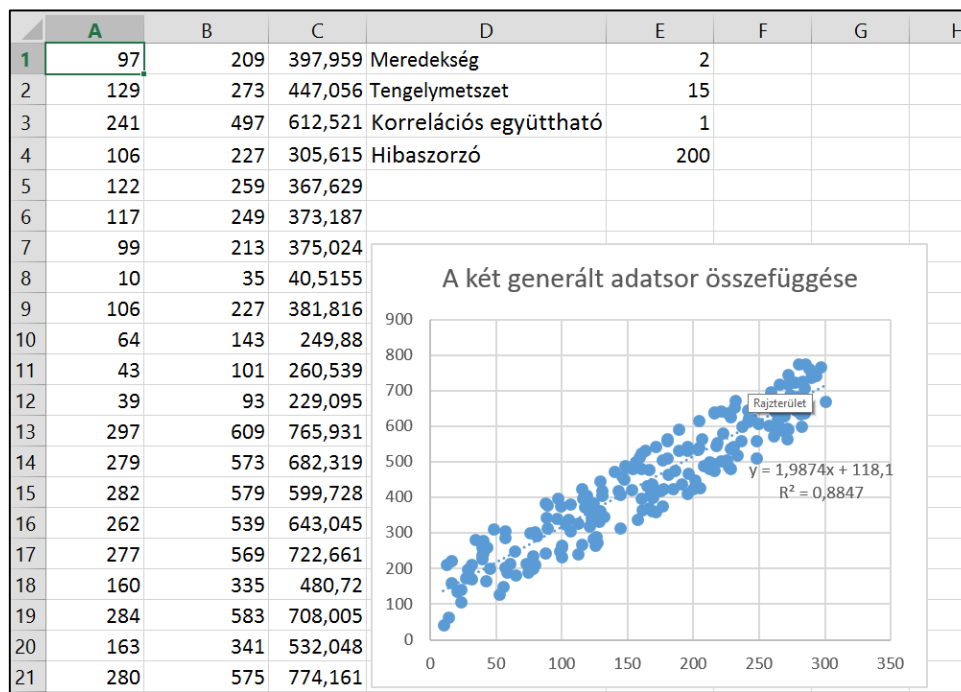
E3 → =KORREL(A1:A200;C1:C200).

Láthatjuk, hogy a korrelációs együttható romlott, de nem „kellő mértékben”, azaz nagyobb hibákra van szükségünk. Amennyiben az E3-as cellában továbbra is 1 szerepel, akkor növeljük meg a tizedeshelyek számát. A hiba mértékét növelhetjük, ha a generált véletlen hibát megszorozzuk egy 1-nél nagyobb számmal. Ehhez az E4-es cellába írunk be egy szorzószámot, például 200-at, majd módosítsuk a C1 cellában lévő függvényt (**32. ábra**):

D4 → Hibaszorzó,

E4 → 200,

C1 → =B1+\$E\$4*VÉL().



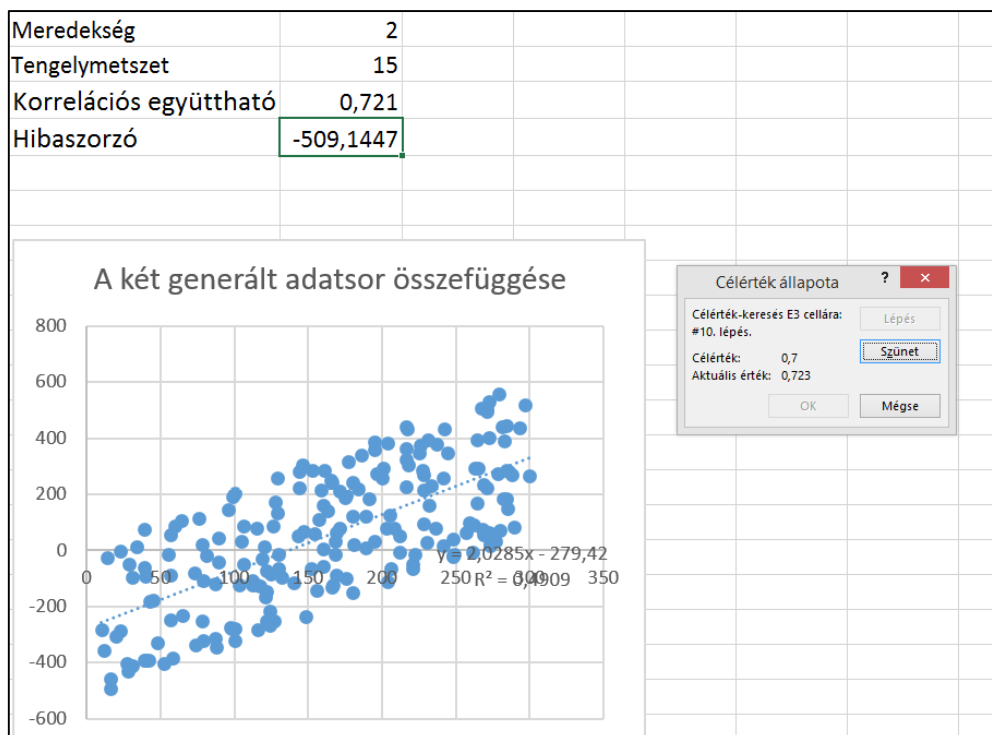
32. ábra: A korreláció alakulása hibaszorzó bevezetésével

Láthatjuk, hogy még mindig nem elég nagy a hiba. A szorzó többszöri módosításával valószínűleg bizonyos idő után megtalálnánk a szükséges értéket, de ilyen esetben, amikor tudjuk, hogy milyen eredményt szeretnénk elérni egy képlettel, de nem ismerjük az ehhez szükséges pontos bemeneti értéket, a célértékkeresés használata javasolt. Ezt a „szolgáltatást” az „Adatok” menüben, az „Adateszközök” csoportban a „Lehetőségelemzés” gombra, majd a „Célértékkeresés” parancsra kattintva érhetjük el egy felugró ablak formájában. Az ablakban 3 adatot kell megadnunk:

- Célcella: azon cellának a hivatkozása, amely a megoldandó képletet tartalmazza (jelen esetben E3)
- Célérték: az elérni kívánt eredmény (jelen esetben 0,7)
- Módosuló cella: a változtatható bemeneti értéket tartalmazó cella hivatkozása (jelen esetben E4)

A szükséges értékek beállítását követően kattintsunk az OK gombra, majd a célértékkeresés a kívánt eredmény eléréséig vagy a maximális iterációs lépésszámig fut. A maximális iterációs számot (alapértelmezésben 100) és az elfogadható hiba mértékét (alapértelmezésben 0.001) a „Fájl” menü „Beállítások” parancsán keresztül a „Képletek” menüpontban állíthatjuk be.

Láthatjuk, hogy az E3-as cella értéke 0,7-hez közeli értéket vett fel az iterációs lépések után (33. ábra). Amennyiben nem értük el a kívánt eredményt a beállított iterációs lépésszámot követően, akkor növeljük meg a maximális iterációs számot!



33. ábra: A korreláció alakulása a célértékkeresés után

Mivel a véletlen számok újragerálódnak a munkalap elemeinek módosítása során, így ezt megelőzve a kívánt eredmény elérését követően a C oszlopot jelöljük ki, másoljuk a vágólapra, és irányítottan, értéként illesszük be a saját helyére.

Összefoglalva, a feladat egy példán bemutatta, hogyan lehet a célértékkeresést alkalmazni egy olyan esetben, amikor tudjuk, hogy milyen eredményt szeretnénk elérni egy képlettel, de nem ismerjük az ehhez szükséges pontos bemeneti értéket.

10. Összefoglaló feladat – a súlyozás problémaköre

Példafájl neve: d-ex.f.xlsx

Adatforrás: a British Antarctic Survey munkatársai, Nemzetközi Atomenergia Ügynökség Víz Erőforrás Programjának GNIP adatbázisa:

<http://www.univie.ac.at/cartography/project/wiser/>

A feladat célja:

Bemutatni a deutériumtöbblet kiszámításának módját alapfüggvényekkel, majd megérteni a csapadékkal, illetve az akkumulációval való súlyozás kérdéskörét. A súlyozás problematikája rendszeresen felmerülő kérdés a természettudományokban, ezért célszerű, ha már a hallgatók a felsőfokú tanulmányuk elején találkoznak ezzel a problémakörrel. Fontos, hogy megértsük a súlyozás lényegét, hiszen nem mindegy, hogy egy adott koncentrációt mekkora mennyiségből határoztunk meg.

Új ismeretek:

- „Hibás” függvénykimenet kezelése
- Hibás kimenet konvertálása

Elméleti háttér:

Stabilizotóp-arány: Az izotóp görög eredetű műszó, az *ἴσος* (*iszosz*, „egyenlő”, „azonos”) és a *τόπος* (*toposz*, „hely”) szavakból képezték. A periódusos rendszer egy elemének nuklidjaira utal („azonos hely”, a periódusos táblázat egy cellája), melyek protonszáma azonos, de neutronjaik száma eltérő, így tömegük is eltér. Stabil izotópoknak azon izotópokat nevezzük, amelyek esetében nem észlelünk radioaktív bomlást. Stabil izotópok esetén hagyományosan nem az abszolút mennyiségeket vagy koncentrációkat mérjük, hanem azok arányait. Ezeket nemzetközi sztenderdekhez, etalonokhoz viszonyított különbséggént fejezzük ki, hogy megvalósulhasson a különböző laboratóriumok eredményei közötti összehasonlíthatóság.

Az oxigénnek három stabil izotópja van, 16-os, 17-es és 18-as tömegszámmal, jelölésük: ^{16}O , ^{17}O , ^{18}O). A 16-os tömegszámú a közönséges (a leggyakoribb), a 18-as tömegszámú a nehezebb, de még mindig viszonylag gyakori (kb. 0,2%), míg a 17-es a legritkábban előforduló izotóp. Amikor a csapadékvíz stabil oxigénizotóp-összetételéről beszélünk, akkor ez alatt jellemzően a legnehezebb és a legkönnyebb izotópok arányát ($^{18}\text{O}/^{16}\text{O}$), illetve annak a nemzetközi sztenderdtől való eltérését értjük (KOHÁN és KERN, 2012).

A számítás módjára az 1-es egyenlet mutat egy példát, ahol az izotópok előfordulásának gyakoriságát a V-SMOW⁵-től való eltérés mértékében, ezrelékben (‰) fejezzük ki (COPLIN, 1994) a „δ” jelölés használatával.

$$\delta^{18}O = \left(\frac{\left(\frac{^{18}O}{^{16}O} \right)_{\text{minta}}}{\left(\frac{^{18}O}{^{16}O} \right)_{\text{standard}}} - 1 \right) * 1000 \text{ ‰} \quad (1)$$

Esetünkben a stabilizotóp-arányokat a Halley Bay antarktiszi állomáson hullott csapadékból mért hidrogén és oxigén stabilizotóp-arányból számított deutériumtöbbslettel (továbbiakban d-többsletként jelöljük és használjuk) fogunk számolni.

Definíció: d-többslet = δD - 8*δ¹⁸O.

A deutériumtöbbslet környezeti jelentősége:

A d-többslet egy származtatott izotóp-változó, (DANSGAARD, 1964). Alkalmas a régmúlt klímájában bekövetkező hirtelen változások nyomon követésére (FÓRIZS, 2003). A d-többslet változásai az alábbi jelenségeket mutathatják: (i) légköri cirkulációban bekövetkezett hirtelen változások, (ii) páraforrás helyén bekövetkezett változások és/vagy (iii) a páraforrás helyének megváltozása (JOHNSON ET AL., 1989), (iv) a csapadék hullási helyén bekövetkezett változások (alapvetően a levegő relatív páratartalma (VODILA ET AL., 2011)).

A d-többslet gyakorlatilag a ¹H¹H¹⁸O molekula ²H¹H¹⁶O-hoz képesti lassabb mozgását tükrözi a diffúzió során. Ennek következménye a HD¹⁶O dúsulása a gáz fázisban pl. felhőképződés során. Ez a különbség azonban csak gyors folyamatok esetén tapasztalható, amikor nincsen idő egyensúlyi állapot kialakulására, pl. párolgás során, amikor a tengerfelszíni relatív páratartalom (<100%), vagy a légáramlatok megakadályozzák az egyensúlyi állapot kialakulását (MERLIVAT és JOUZEL, 1979). Bizonyított, hogy a tengerfelszíni hőmérséklet és a határrétegben a páratartalom jelentős szerepet játszik a d-többslet alakulásában (STEIG, 2003), azonban ezek relatív mértéke továbbra is vita tárgyát képezi (PFAHL és SODEMANN, 2014).

Az itt bemutatott feladat kettős célt tűz maga elé: (i) a d-többslet kiszámítása csapadékból nyert stabilizotóp-arányok alapján és (ii) a súlyozás általános problémakörének körbejárása. Nem mindegy, hogy egy adott izotóp-koncentrációt mekkora csapadékmennyiségből, a vizsgált területen az Antarktiszon lehulló hóból (precipitation), vagy már a talajon megtalálható hórétegből (accumulation) mérték meg.

A súlyozást az alábbi képlet szerint kell elvégezni minden évre. Ideális esetben, amikor nincs adathiány (2), az alábbi egyenlettel kell számolnunk:

⁵ Vienna Standard Mean Ocean Water, a bécsi Nemzetközi Atomenergia Ügynökség tengervíz referenciamintája

$$d - \text{többllet súlyozott} = d.ex_i * \frac{prec_i}{prec.ann} \quad (2)$$

Magyarázat: $i = 1-12$ -ig terjed, a hónapok sorszáma, $d.ex$ a d -többllet, $prec.$ a csapadék (hó) mennyisége, $prec.ann$ pedig az éves csapadékösszeg.

Abban az esetben, ha bármelyik izotópot (H vagy O) nem mérték meg a csapadékból, már nem tudjuk minden csapadékeseménynél kiszámolni a d -többlletet, így ezen csapadékeseményeket ki kell vonnunk a teljes éves átlagból (3) az alábbiak szerint:

$$d - \text{többllet súlyozott} = d.ex_i * \frac{prec.i}{prec.ann - prec.no.d} \quad (3)$$

Itt $prec.no.d$ azon csapadékesemények összege az adott évre, melyekhez nem tartozott d -többllet mérés.

Más szavakkal, az adott hónaphoz tartozó d -többlletet megszorozzuk a hozzá tartozó csapadék mennyiségével, majd elosztjuk az éves csapadékösszeggel, amelyből ki kell vonni azon csapadéértékeket, amelyekből nem lett a d -többllet méréséhez szükséges izotóp mintázva (Hatvani és Kern, 2016).

Feladatok:

- (a) Számítsuk ki a d -többlletet ($d.ex$)
- (b) Súlyozzuk a kapott d -többllet értékeit

Megoldás:

(a) feladat:

Nyissuk meg a ***d-ex.f.xlsx***-et! Az „Alapadat” munkalapon az oxigén- (C oszlop) és hidrogénizotóp (D oszlop) értékeket találjuk. Figyelembe kell venni, hogy mérték-e mindkét paramétert. Ha nem, akkor ugyanis téves eredményt kapunk. Ezért egy kétszeresen egymásba ágyazott HA függvénnyel megvizsgáljuk, hogy teljesül-e a szükséges feltétel. Amennyiben, akár a deutérium, akár az oxigén esetében hiányzik a mért érték, akkor a $d.ex$ -t nem számíthatjuk ki, ezért ilyenkor a „NA”-t íratunk ki.

Az „Alapadat” munkalapon lévő értékeket átmásoljuk a „D-többllet kiszámítása példa” munkalapra, ahol az alábbiakat végezzük el:

$E2 \rightarrow =HA(C2="" ; "NA" ; HA(D2="" ; "NA" ; C2-8*D2))$.

A képletet lemásoljuk az utolsó vonatkozó sorig, (esetünkben a 37. sorig) **(34. ábra)**.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	dátum	IAEA_Precipitation	H2	O18	d-excess								
2	1973.01.15	0.9	-83	-11.13	=HA(C2="";"NA";HA(D2="";"NA";C2-8*D2))								
3	1973.02.15	10	-86.2	-11.96	HA(logikai_vizsgálat; [érték_ha_igaz]; [érték_ha_hamis])								
4	1973.03.15	18.5	-116.8	-15.49	7.12								
5	1973.04.15	11.2	-141.7	-16.89	-6.58								
6	1973.05.15	-0.6	-185.8	-24.09	6.92								
7	1973.06.15	7.4	-223	-28.84	7.72								
8	1973.07.15	-0.2	-134.1	-16.7	-0.5								
9	1973.08.15	0.01	-189.4	-25.73	16.44								
10	1973.09.15	6.3	-185	-24.98	14.84								
11	1973.10.15	8.7	-155	-19.59	1.72								
12	1973.11.15	6.3	-131	-17.82	11.56								
13	1973.12.15	4.7	-113.7	-13.47	-5.94								
14	1974.01.15	9.2	-95	NA									
15	1974.02.15			-12.46	NA								
16	1974.03.15	15.3		-18.59	NA								
17	1974.04.15	1.6		-22.97	NA								
18	1974.05.15	0.3		-21.63	NA								

34. ábra: A d-többlet kiszámítása

(b) feladat:

A kapott E oszlopot irányítottan, értékként bemásoljuk a „Munkaanyag” munkalap B oszlopába, és mellé másoljuk a csapadékot a C oszlopba.

Mivel kulcskérdés a rendelkezésre álló adatok száma, ezért meg kell vizsgálni, hogy egy adott évben hány alkalommal mérték a csapadékot. Amennyiben kevesebb, mint az esetek – vagyis a hónapok – felében volt csak csapadékmérés, úgy nem lehet megbízhatónak tekinteni a súlyozást, így az adott évet ki kell hagyni.

Most a „Munkaanyag” lapon DARAB függvénnyel megszámoljuk a mért csapadéktételek számát, minden adott évre egy külön cellában, az alábbiak szerint:

D2 → =DARAB(C2:C13).

Értéke 12 lesz, vagyis az első esztendőben, 1973-ban 12 hónapban mérték csapadékot.

Most ki kell jelölni a D2:D13 cellatartományt, kimásolni CTRL+C-vel és beilleszteni D14:D37 tartományba, akárcsak az 4. fejezet „(e)” részfeladatban. (Eszerint 1974-ben 11, 1975-ben 4 hónapban mérték csapadékot.)

Jól látható, hogy a harmadik vizsgált évben (1975) nem áll rendelkezésre elegendő számú csapadékadat a súlyozáshoz, így azt az évet a további számításokból ki kell majd hagyni. Ezt az információt a későbbiekben fogjuk használni.

Mivel voltak olyan esetek, amikor nem volt elegendő adat a d-többlet kiszámítására, ezeket az „NA” szöveggel jelöltük meg, így a megfelelő időpontokban mért csapadékot sem szabad felhasználni, azt ki kell vonni a teljes csapadékösszegekből. Más szavakkal, ahol a „d-excess” (B) oszlopában NA jelzés van, ott a „Precipitation” (C) oszlop értékét „meg kell

tartani”, ahol pedig nincs NA, ott a C oszlop értéke helyett 0-t kell használni. Ezt FKERES (Megoldás 1; F oszlop), vagy HA függvényekkel (Megoldás 2; G oszlop) tehetjük meg. Írjuk be az alábbi képletet az E2 cellába és másoljuk le a 37. sorig:

E2 → =FKERES("NA";B2:C2;2;HAMIS).

Az FKERES esetében a „#HIÁNYZIK” kimenetet kapjuk minden olyan esetre, ahol hiányzik a csapadékmérés. Ezeket át kell konvertálnunk nullává, hogy összegezhessük azon csapadéértékeket, melyekhez nem tartozott számított d-többlet („d.excess”). Ezt az alábbi képlettel tehetjük meg:

F2 → =HA(HIBÁS(FKERES("NA";B2:C2;2;HAMIS));0;E2).

Ennél egy egyszerűbb megoldás a nem létező d-többlethez tartozó csapadéértékek kigyűjtéséhez a már jól ismert HA függvény használata:

G2 → =HA(B2="NA";C2;0).

Mind a két megoldás végeredményként ugyanazt az adatsort kapjuk.

Következőben összegezzük azon csapadéértékeket, melyekhez nem tartozik d-többlet:

H2 → =SZUM(F2:F13).

Miután a H2 cellába bevittük az előző képletet, a korábbi gyakorlatnak megfelelően lemásoljuk. Kijelöljük a H2:H13 tartományt, és lemásoljuk egész számú többszörös tartományába H14:H37-ig.

Miután előkészítettünk minden peremfeltételt, megnéztük, hogy van-e megfelelő számú csapadékadat, amivel súlyozni lehet, illetve kigyűjtöttük és összegeztük azon csapadéértékeket, melyekhez nem tartozik d-többlet, elvégezhetjük a súlyozást az I oszlopban. Itt újabb feltételeket kell figyelembe venni:

- ha nincs d-többlet adat, „NA”-t adjon, ki, de amennyiben van:
 - vizsgálja meg, hogy van-e a vizsgált időtartam felénél több csapadék-adat, ha nincs,
 - akkor adja meg, hogy kevés az adat („KA”),
- különben pedig végezze el a súlyozást az alábbiak szerint:

Az adott d-többlet értékét szorozzuk meg a hozzá tartozó csapadékkal, majd osszuk el az éves csapadékösszeggel, amelyből ki kell vonni azon csapadéértékeket, amelyekből a d-többletet nem mintázták.

A számolást éves bontásban tudjuk elvégezni. A fő probléma, hogy nincs az Excelben olyan függvény, ami egy feltételtől függően változtatná a cellahivatkozások rögzítését. Pedig pontosan ez az, amire szükségünk lenne, hogy „elengedjük” a D\$2; H\$2 és C\$2:C\$13 -as cellák rögzítését, és évente 12 cellával továbblépjünk, 1974-ben pl. a D\$14; F\$14 és C\$14:C\$25 cellákba.

Manuálisan azonban kétféleképpen is át tudjuk hidalni ezt a problémát.

1. lehetőség:

Létre kell hozni az első évet, mint mintaévet, és annak a függvényhivatkozásait lemásolni egész számú többszörösként. Első lépésben az alábbiak szerint:

$I2 \rightarrow =HA(B2="NA";B2;HA(D2<6;"KA";B2*C2/(SZUM(C2:C13)-H2))))$,

$I3 \rightarrow =HA(B3="NA";B3;HA(D2<6;"KA";B3*C3/(SZUM(C2:C13)-H2))))$,

és így tovább egészen az év végéig:

$I13 \rightarrow =HA(B13="NA";B13;HA(D2<6;"KA";B13*C13/(SZUM(C2:C13)-H2))))$.

Miután létrehoztuk a teljes év hivatkozásait és kiszámoltuk a megfelelő d-többleteket, jelöljük ki a teljes évet I2:I13, másoljuk ki és illesszük be a cellák számának egész számú többszörösének megfelelő cellatartományba (I2:I37).

Jól látszik, hogy a képletünk működik, hiszen 1974-ben mindösszesen két esetben számolt d-többletet, mivel nem is volt több mért értékünk, míg 1975-re azt írta ki, hogy kevés az adat (KA), mivel az 1975-ben mért csapadékadatok száma kevesebb volt, mint 6 hónap (6 darab, azaz 50%).

2. lehetőség:

Továbbra is beírjuk az eredeti képletet az I2 cellába:

$I2 \rightarrow =HA(B2="NA";B2;HA(D2<6;"KA";B2*C2/(SZUM(C2:C13)-H2))))$.

Másoljuk át I2-t I3-ba, majd csökkentsük a kapott D3, C3:C14, H3 indexeit 1-gyel és dollárjellel rögzítsük: D\$2, C\$2:C\$13, H\$2-ként:

$I3 \rightarrow =HA(B2="NA";B2;HA(D$2<6;"KA";B2*C2/(SZUM(C$2:C$13)-H$2))))$.

Ezután másoljuk I3-at I4:I13-ba. Ezzel az első év elkészült, és ehhez 1 képletet (az I2-ét) be kellett írni, 4+4=8 karaktermódosítást és 2 másolást kellett végezni. Ugyanezt a lépéssort kell elvégezni minden évre elsőként az I2-ben lévő képlet átmásolásával.

Összegezve: megtanultuk, hogy hogyan lehet felépíteni az eddig megtanult képletekből, funkciókból egy többlépcsős folyamatot, továbbá, hogy a hibás kimenetet („#HIÁNYZIK”), hogyan tudjuk számszerűsíteni és felhasználni egy probléma megoldása során.

Köszönetnyilvánítás

A szerzők igazán hálásak az ELTE TTK Általános és Alkalmazott Földtani Tanszéke munkatársainak a diszkusszióért. A jegyzet létrejöttét jelentősen segítették azok a föld- és környezettudomány (B.Sc.) szakos hallgatók, akik ezen a példák egy részét az oktatás során megismerték, és a szerzőkkel megosztották tapasztalataikat.

Hatvani István Gábor munkája a TÁMOP 4.2.4.A/1-11-1-2012-0001 azonosítójával „Nemzeti Kiválóság Program – Hazai hallgatói, illetve kutatói személyi támogatást biztosító rendszer kidolgozása és működtetése országos program” című kiemelt projekt keretében zajlott. A projekt az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósult meg. Köszönet illeti továbbá a az MTA „Lendület” programját (LP2012-27/2012) is. A jegyzet elkészítését támogatta a MOL Nyrt., az ELTE-vel kötött „MOL Mecenatúra szerződés” keretében.

Felhasznált irodalom

- Bácsatyai, L. 2005. Magyarországi vetületek, Hivatkozva: 2015.07.28
http://www.geo.info.hu/geodezia/dokumentumok/geod-vettan/magyarorszgi_vetletek.pdf
- Box, G.E.P., Jenkins, G.M., Reinsel, G.C. 1994. Time Series Analysis: Forecasting and Control. Englewood Cliffs, NJ: Prentice Hall.
- Coplen, T.B. 1994. Reporting of stable hydrogen, carbon and oxygen isotopic abundances. Pure App. Chem. 66.273-76.
- Dansgaard, W. 1964. Stable isotopes in precipitation. Tellus 16, 436-68.
- EC 2000. Directive 2000/60/EC of the European Parliament and of the Council establishing a framework for Community action in the field of water policy. OJ L 327, 22.12.2000
- Fórizs, I. 2003. Isotopes as natural tracers in the water cycle: Examples from the Carpathian Basin. Studia Universitatis Babeş-Bolyai, Physica, Special Issue 48, 69-77
- Haagen-Smit, A.J. 1952. "Chemistry and physiology of Los Angeles smog." *Industrial & Engineering Chemistry* 44(6), 1342-1346.
- Hatvani, I. G., Magyar, N., Zessner, M., Kovács, J., & Blaschke, A. P. 2014. The Water Framework Directive: Can more information be extracted from groundwater data? A case study of Seewinkel, Burgenland, eastern Austria. *Hydrogeology Journal*, 22(4), 779-794.
- Hatvani, I. G. és Kern, Z. 2016. Weighting alternatives (precipitation vs. accumulation) for water stable isotopes – implications for ice core records. *Front. Earth Sci.* 4(20), doi: 10.3389/feart.2016.00020
- Johnsen, S. J., Dansgaard, W., és White, J. W. C. 1989. The origin of Arctic precipitation underpresent and glacial conditions, Tellus B, 41, 452–468.
- Kohán, B. és Kern, Z. 2012. A felszíni csapadék stabil oxigénizotóp összetételének becslése Magyarországon térinformatikai módszerekkel - előzetes eredmények. In: Lóki, J. (ed.) AZ ELMÉLET ÉS A GYAKORLAT TALÁLKOZÁSA A TÉRINFORMATIKÁBAN III. Debreceni Egyetemi Kiadó, Debrecen pp. 189-196.
- Kovács, J., Korponai, J., Kovácsné Székely, I., Hatvani, I.G. 2012a Introducing sampling frequency estimation using variograms in water research with the example of nutrient loads in the Kis-Balaton Water Protection System (W Hungary). *Ecological Engineering* 42, 237-243.
- Kovács, J., Nagy, M., Czauner, B., Székely Kovács, I., Kériné Borsodi, A., Hatvani, I.G. 2012b Delimiting sub-areas in water bodies using multivariate data analysis on the example of Lake Balaton (W Hungary), *Journal of Environmental Management* 110, 151-158.
- Kovács, J., Tanos, P., Korponai, J., Kovácsné, Sz. I., Gondár, K., Gondár-Sőregi, K., Hatvani, I.G. 2012c Analysis of Water Quality Data for Scientists In: Kostas Voudouris; Dimitra Voutsas (eds.) *Water Quality and Water Pollution: Evaluation of Water Quality Data*. Rijeka: InTech Open Access Publisher
- Kovalcsik, G. 2010. Az Excel programozása. COMPUTERBOOKS Budapest, 306p.

- Liberty, J. 2003. Programming Visual Basic.NET 2nd Edition. O'Reilly Media 560p.
- Magyar, N., Hatvani, I. G., Székely, I. K., Herzig, A., Dinka, M., & Kovács, J. 2013a. Application of multivariate statistical methods in determining spatial changes in water quality in the Austrian part of Neusiedler See. *Ecological Engineering*, 55, 82-92.
- Magyar, N., Trásy, B., Kutrucz, Gy., Dinka, M. 2013b. Delineating water bodies on the Hungarian side of Lake Fertő/Neusiedler See. In: Elemér Pál-Molnár (szerk.) : Theories and applications in geomathematics. Szeged: GeoLitera. pp. 103-115.
- Merlivat, L. és Jouzel, J. 1979. Global climatic interpretation of the deuterium-oxygen 18 relationship for precipitation, *J. Geophys. Res.*, 84, 5029–5033,
- Microsoft 2013. Microsoft Excel [számítógépes program és online súgó]. Redmond, Washington: Microsoft. Hivatkozva: 2015.08.18. <https://support.office.com/#>
- PAC 1990, 62, 2167. Glossary of atmospheric chemistry terms (Recommendations 1990), doi: 10.1351/pac199062112167
- Pfahl, S. és H. Sodemann. 2014. What controls deuterium excess in global precipitation? *Clim. Past* 10, 771-81.
- Pongrácz, R., Levegőszennyezés és éghajlaváltozás In: Szabó, M. és Angyal, Zs. (szerk.) 2012. A környezetvédelem alapjai, Eötvös Loránd Tudományegyetem Természettudományi Kar. Typotex Kiadó 341 p. Hivatkozva: 2015.07.28. http://www.tankonyvtar.hu/hu/tartalom/tamop425/0047_Angyal-Szabo_Kornyezetvedelem/A_kornyezetvedelem_alapjai_OK.pdf
- Salma, I., Légkörkémi In: Salma I. (szerk.) 2012. Környezatkémia, Eötvös Loránd Tudományegyetem Természettudományi Kar. Typotex Kiadó 341 p. Hivatkozva: 2015.07.28. http://www.tankonyvtar.hu/hu/tartalom/tamop425/0047_Salma_Kornyezetkemia/Salma_kornyezetkemia.pdf
- Steig, E.J. 2003. Ice cores. *Paleoclimatology*, 1673-1680
- Tatár, E. és Óvári, M. (szerk.) 2012. Környezeti mintavételezés, Eötvös Loránd Tudományegyetem Természettudományi Kar. Typotex Kiadó 150 p. Hivatkozva: 2015.07.28. http://www.tankonyvtar.hu/hu/tartalom/tamop425/0047_Ovari_Kornyezeti_mintavetelezes/Ovari_Kornyezeti_mintavetel.pdf
- Vodila, G., Palcsu, L., Futo, I., Szanto, Z. 2011. A 9-year record of stable isotope ratios of precipitation in Eastern Hungary: Implications on isotope hydrology and regional palaeoclimatology *JOURNAL OF HYDROLOGY* 400:(1-2), 144-153.