

## Egy felnőttképzési MOOC-programon részt vevők aktivitási adatainak elemzése: lemorzsolódás, sikeresség és tanulási profilok

A tömeges nyílt online kurzusok (*Massive Open Online Courses, MOOC*) olyan távoktatási képzések, amelyek a résztvevők nagy számával és heterogenitásával, valamint a magas lemorzsolódási aránnyal tűnnek ki. Mindezek megnehezítik a résztvevők előrehaladásának nyomon követését. Jelen kutatás során egy felnőttképzési környezetben megvalósított, tömeges nyílt online kurzus résztvevőinek ( $N = 7760$ ) tanulói profiljait vizsgáltuk. A tanulói profilok meghatározását a résztvevők tanulási aktivitásának alapján végeztük el, amelyhez látens profil (*Latent Profile Analysis, LPA*) elemzést használtunk. A résztvevők aktivitása, kurzus iránti elkötelezettsége a kurzus első felében jelentős mértékben, a második felében csak kis mértékben csökkent. A kurzus záróvizsgáján való sikerességet legnagyobb mértékben a tananyagban való előrehaladás határozta meg. A résztvevők aktivitása alapján kilenc tanulói profilt azonosítottunk, amelyek két nagyobb kategóriába sorolhatók. A tanulói profilok egy része esetén a záróvizsgán nyújtott teljesítmény arányos volt a résztvevők aktivitásával, más részüknél nem állt fenn ez az összefüggés. A tanulói profilok azonosítása megalapozhatja a személyre szabott oktatás lehetőségét.

**Kulcsszavak:** tanulói aktivitás, MOOC, klaszteranalízis, látens profil elemzés, tanulói profilok

### Köszönetnyilvánítás

A tanulmány az Innovációs és Technológiai Minisztérium Kooperatív Doktori Program Doktori Hallgatói Ösztöndíj Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült. A kutatást az OTKA K135727, illetve az MTA Közoktatás-fejlesztési Kutatási Pályázata támogatta (KOZOKT2021-16).

### Szerzői információ

**Csányi Róbert**, SZTE Neveléstudományi Doktori Iskola, <https://orcid.org/0000-0003-1740-8582>  
**Lőrincz Mónika Mária**, Nemzeti Szakképzési és Felnőttképzési Hivatal  
<https://m2.mtmt.hu/gui2/?type=authors&mode=browse&sel=10075320>  
**Molnár Gyöngyvér**, SZTE Neveléstudományi Intézet, MTA-SZTE Digitális Tanulási Technológiák Kutatócsoport, <https://orcid.org/0000-0003-4890-6904>

### Így hivatkozzon erre a cikkre:

Csányi Róbert, Lőrincz Mónika Mária, Molnár Gyöngyvér. „Egy felnőttképzési MOOC-programon részt vevők aktivitási adatainak elemzése”. *Információs Társadalom* XXIV, 2. szám (2024): 34–55.

<https://dx.doi.org/10.22503/inftars.XXIV.2024.1.2>

A folyóiratban közölt művek  
a Creative Commons Nevezd meg! – Ne add el! – Így add tovább! 4.0  
Nemzetközi Licenc feltételeinek megfelelően használhatók.

## Learning profiles of participants in an adult education MOOC program

Massive Open Online Courses (MOOCs) are distance education courses that stand out for their large number and heterogeneity of participants and high drop-out rates. This makes it difficult to monitor the progress of participants. In the present study, we investigated the learning profiles of participants (N = 7760) in a massive open online course in an adult learning environment. Learner profiles were determined based on the learning activity of the participants using Latent Profile Analysis (LPA). Participants' activity and engagement in the course decreased significantly in the first half of the course and only slightly in the second half. Success in the final exam was most strongly determined by progress in the course material. Based on the activity of the participants, nine learning profiles were identified, which can be grouped into two broad categories. The identification of student profiles can provide a basis for personalised learning.

**Keywords:** *student activity, MOOC, cluster analysis, latent profile analysis, student profiles*

---

## Bevezetés

Az informatikai eszközök 20. század végi és 21. századi jelentős fejlődésével és széles körű elérhetővé válásával lehetőség nyílt a technológia tanítási-tanulási folyamatban való mindennapos használatára. A technológia használata a teljes pedagógiai folyamatot képes felölelni, kiterjed az oktatás minden szintjére az óvodától a felsőoktatáson át a felnőttkori tanuláshoz. A technológia oktatásban való használatának egyik legnagyobb előnye, hogy tovább lehet lépni a hagyományos, frontális, tanárközpontú „one size fits all” szemléleten, és elérhetővé válhat a személyre szabott tanulás lehetősége (Molnár 2021).

A technológia oktatásban való alkalmazása nem csak napjaink igénye, már a digitális korszakot megelőzően, a 20. század első felében megjelentek az első mechanikus oktatógépek. Az első oktatógépet Sidney L. Pressey fejlesztette ki az 1920-as években azzal a megfontolással, hogy az oktatásban számos olyan feladatot kell a tanulókkal rutinszerűen gyakoroltatni, amely feleslegesen nagy terhet jelent az oktatók számára. Ekkor a tanulók még papíron kapták meg a számozott kérdéseket a négy válaszlehetőséggel együtt, és a gépen a négy válaszlehetőség nyomógombjai közül kellett a megfelelőt kiválasztani. Amennyiben egy tanuló jó választ adott, akkor a gép továbblépett a következő feladatra, ha nem, akkor a tanuló tovább próbálkozhatott, ezáltal a tanulók a nekik megfelelő tempóban haladhattak előre a feladatokban (Kassymova et al. 2020). Az oktatógépek ezzel megtették az első lépést a személyre szabott oktatás felé, ugyanakkor annak valódi lehetősége napjainkban vált realitássá.

A modern technológiák oktatásban való használata során olyan kontextuális adatokat (feladattal eltöltött idő, előre- és hátraugrások, szemmozgás, kattintások stb.) lehet gyűjteni, amelyek a hagyományos oktatási és értékelési rendszerekben elképzelhetetlenek (Tóth et al. 2017). A kontextuális adatok elemzése alapján még inkább személyre lehet szabni az oktatást, amivel tovább növelhető az oktatás hatékonysága. Az oktatók objektív visszajelzéseket kaphatnak a tanítási-tanulási folyamatról, amely alapján megvalósítható a tanulók tanulásának és viselkedésének elemzése (Romero és Ventura 2020).

### *Tömeges nyílt online kurzusok*

A tömeges nyílt online kurzus (Massive Open Online Course, MOOC) kifejezést David Cormier alkotta meg egy online kurzusra vonatkozóan, amelyet a kanadai Manitoba Egyetemen hirdettek meg 2008 őszi félévében (Hollands és Tirthali 2014). A kifejezést érdemes szóról szóra elemezni. A *tömeges* azt jelenti, hogy a résztvevők száma sokkal nagyobb, mint az a hagyományos oktatás keretében elképzelhető. A *nyílt* kifejezés az ingyenes (vagy alacsony áron elérhető) hozzáférhetőséget és a bárki számára való nyitottságot jelenti, ugyanis a kurzusok többségének nincs semmilyen belépési követelménye. A kurzusok az *online* térben zajlanak, az internet segítségével a nap bármely szakában bárhol is hozzáférhetőek. A *kurzus* kifejezés pedig a strukturált tananyagra utal, amelyek által a résztvevők önállóan is képesek tanul-

ni (Khalil és Ebner 2017). A MOOC-ok lehetőséget teremtenek, hogy bárki *akármit, akárhol, akármikor* tanulhasson (Longstaff 2017), mindezt úgy, hogy „bárki számára ingyen elérhetővé tették az élvonalbeli egyetemek szupersztár professzorainak kurzusait” (Molnár, Turcsányi-Szabó és Kárpáti 2020).

A MOOC-ok általában független platformszolgáltatókon keresztül érhetőek el, közülük a legismertebbek a Coursera, az edX és az Udacity (Hollands és Tirthali 2014). A hagyományos oktatási formákhoz képest jelentős különbség, hogy itt sokkal nagyobb a résztvevők heterogenitása, amely többek között az iskolai végzettségben, az előzetes ismeretek mértékében, az életkorban, a szociokulturális háttérben is megnyilvánul (Barthakur et al. 2021; Ortega-Arranz et al. 2019). Az első MOOC-okat egyetemek indították, így a korai HarvardX és MITx kurzusokon a medián életkor 30 év alatt volt (Ho et al. 2014). A MOOC-ok szélesebb körben való ismertté válásával és elterjedésével a résztvevők átlagéletkora is magasabb lett. Williams és munkatársai (2018) nyolc edX kurzus, összesen több mint 15 000 résztvevőjének életkorát hasonlították össze, az átlagéletkor az egyes kurzusokon 32 és 41 év között volt. További fontos különbség, hogy a résztvevőknek a folyamatos tanári jelenlét és instrukciók helyett videók, majd azt követő kérdések és feladatok segítségével alapvetően önállóan kell tanulniuk (Yu, Wu és Liu 2019).

A MOOC-oknak a hagyományos oktatási formákhoz képest számos előnyük van. Sokkal változatosabb képzési kínálat nyújtható belőlük, így a tanulók az igényeiknek, preferenciáiknak és helyzetüknek megfelelő képzésekre jelentkezhetnek (Matcha et al. 2020). Könnyen hozzáférhetőek és az egyéni időbeosztáshoz igazítható a kurzusok elvégzése (Khalil és Ebner 2017). Alacsony költséggel járnak, mivel nem kell elutazni a képzésekre, nagy létszámmal indíthatók, így alacsony az egy résztvevőre jutó költség, valamint egy kurzus tananyagát egyszer kell elkészíteni, ugyanakkor többször is felhasználható, többször is indítható és kiközvetíthető (Námesztovszki et al. 2015).

A MOOC-ok ugyanakkor számos előnyük ellenére, sőt részben épp a fent említett előnyeikből fakadóan jelentősen nagyobb lemorzsolódási rátával valósulnak meg, mint a jelenléti kurzusok. Egy-egy kurzusra nagyon sok jelentkező regisztrál, ugyanakkor közülük sokan el sem kezdik a képzést, valamint a kurzusokba belekezdők közül is sokan lemorzsolódnak, nem fejezik be a képzést (Khalil és Ebner 2017). Hollands és Tirthali (2014) elemzésében 85% és 97% között megvalósuló lemorzsolódási arányt azonosított, míg mások átlagosan 90% körülit tapasztaltak (Gregori et al. 2018; Hew és Cheung 2014; Morris 2013). A magas lemorzsolódási arány oka egyrészt a résztvevők oldaláról (1) a motiváció hiánya, (2) az időhiány és (3) a nem megfelelő előzetes tudás, másrészt a kurzus jellegéből adódóan (4) a kurzus esetlegesen nem megfelelő kialakítása, (5) az elszigeteltség, (6) az alacsony szintű interaktivitás és (7) az oktatókkal való kapcsolat hiánya, ami személytelenné teszi a tanítást és tanulás folyamatát és sok résztvevőre negatívan hat (Khalil és Ebner 2017; Morris 2013; Zhu et al. 2022). A magas lemorzsolódási arány egy olyan probléma, amellyel minden MOOC-platform küzd, csökkentését különböző felépítésű, szerkezetű, motivációs elemeket tartalmazó kurzusok és vonatkozó kutatások indításával és alkalmazásával igyekeznek megoldani (Yu, Wu és Liu 2019).

A MOOC-platformokon kínált kurzusok száma évről évre nő, a Class Central adatai alapján 2021-ben a világ több mint 900 egyeteme közel 20 000 kurzust kínált

---

(Class Central 2021). A MOOC-ok növekvő népszerűségének legfőbb oka Majó-Petri és munkatársai (2020) szerint nem a nyílt hozzáférés vagy az ingyenesség, hanem a fiatalabb generációk változó tanulási szokásai, valamint a tömegesen megvalósuló, személyre szabott oktatás lehetősége. A hallgatók nagy létszáma miatt a kurzusok során nagy mennyiségű adat keletkezik, amelyek elemzése alapján jobb minőségű tananyagok állíthatók elő, amelyek így egyre inkább megfelelnek a hallgatók személyre szabott igényeinek (Szani és Merkovity 2014).

### *Tanulói profilok meghatározása*

A MOOC-kutatások egyik fő célja a diákok online környezetben való tanulásának minél jobb megértése, a hasonló tulajdonságokkal rendelkező tanulók csoportjainak azonosítása. Ehhez leggyakrabban azt elemzik, hogy a résztvevők hogyan használják a kurzus keretein belül elérhető különböző eszközöket és forrásokat (videók, kvízek, chat stb.). A tanulói csoportok azonosítására a korábbi kutatások leggyakrabban k-közép vagy hierarchikus klaszteranalízist használtak (Barthakur et al. 2021; Del Valle és Duffy 2009; Fan et al. 2022; Rodrigues et al. 2016). A klaszteranalízis alkalmazásával a hétköznapi életben is széleskörűen találkozunk, ez alapján szegmentálják a vevőket a webáruházak, így működnek a streamingszolgáltatók ajánlói, vagy éppen ilyen módszerekkel próbálják megelőzni az internetes csalásokat (Henning et al. 2015).

Rodrigues és munkatársai (2016) egy 5100 fős brazil MOOC résztvevőinek csoportosítását végezték el egyrészt hierarchikus (Ward), másrészt nem hierarchikus klaszteranalízissel (k-közép). A résztvevők hét csoportját különítették el, amely csoportok összességében háromféle elköteleződési mintázatot tükröztek. A résztvevők 16%-a bizonyult nagy mértékben elkötelezettnek a kurzus teljesítése iránt, 26%-uk részlegesen, míg 58%-uk nem volt elkötelezett. Del Valle és Duffy (2009) egy online tanártovábbképzési kurzus 59 résztvevőjének viselkedését elemezte hierarchikus (Ward) klaszteranalízissel. Nyolc változót vontak be az elemzésbe, amelyek alapján szintén három klaszter alakult ki. Az első klaszterbe tartozott a tanárok nagyobb része (59%), ők töltötték a legtöbb időt a feladatokkal és a legtöbb tanulási forrást használták fel, ők a tananyag teljes megtanulására törekedtek. A második klaszterbe tartozók (22%) gyakran és rendszeresen megoldották a rendszer által kiközvetített feladatokat, de jóval kevesebb időt töltöttek a tananyag feldolgozásával. A harmadik klaszterbe tartozók (19%) összességében a másik két csoporthoz képest a legkevesebb idő alatt minimális erőfeszítést tettek a kurzus elvégzése érdekében.

A látens csoport elemzés (*Latent Class Analysis, LCA*) egy olyan statisztikai alapon nyugvó klaszterelemzési technika, amelyet gyakran használnak az emberi viselkedésminták profilozására és megértésére. A hagyományos klaszterezéssel egy adott tanuló egy adott tulajdonsággal jellemezhető csoportba tartozik. A látens profil elemzés valószínűségi alapon nyugszik, emiatt egy adott tanuló több csoportba is tartozhat. Meg tudjuk mondani, hogy egy tanuló mekkora valószínűséggel tartozik egy adott tanulási stratégiát használó csoportba, és a gyakorlatban abba a csoportba soroljuk, ahova a legnagyobb valószínűséggel tartozik (Henning et al. 2015).

A látens csoport elemzést eredetileg kategorikus és ordinális változók elemzésére használták, később azonban egy módosított algoritmus segítségével megvalósult ezen elemzési eljárás alkalmazása folytonos változók esetén is. Ebben az esetben a módszert gyakran látens profil elemzésnek (*Latent Profile Analysis, LPA*) nevezik (Muthén és Muthén 2012). Barthakur és munkatársai (2021) egy négy kurzusból álló MOOC-program 175 résztvevőjének kontextuális adatait elemezték. A látens profil elemzésbe 23 változót vontak be, amivel hatféle tanulási stratégiát azonosítottak. (1) Intenzív (14,5%): magas szintű elkötelezettség és az összes tanulási forrás használata. (2) Értékelésorientált (25,4%): a legszélesebb körben használt stratégia, amelyet elsősorban az összes értékelési feladatban való magas szintű aktivitás és a kvízkérdések megoldása jellemez. (3) Nagymértékben demotivált (21,2%): az ide tartozók voltak a legkevésbé elkötelezettek a kurzus iránt, ők a többiekénél jelentősen alacsonyabb pontszámot értek el az összes változó esetében. (4) Videófókuszú (1,5%): elsődleges jellemzői a videós tevékenységekben való fokozott aktivitás és a rendkívül intenzív videóéletöltés. (5) Mérsékelt elkötelezettség (13,4%): ezt a stratégiát a kurzus átlagos használata jellemezte. (6) Demotivált (23,9%): ez a stratégia minimális aktivitást tükröz az összes rendelkezésre álló komponens és erőforrás tekintetében. Összefoglalva kiemelhetjük, hogy a klaszterek száma, valamint a létrejövő klaszterek jellemzői erősen függenek számos tényezőtől, többek között magától a kurzustól, az elemzésbe bevont változóktól és a klaszterezéshez használt módszertől.

### *Célok, kutatási kérdések*

A nemzetközi szakirodalomban széles körben vizsgált téma a MOOC-résztvevők profiljainak azonosítása, azonban legjobb tudomásunk szerint limitáltak a hazai vonatkozó kutatások. A jelen kutatás célja a résztvevők MOOC-kurzus során detektált aktivitásának, valamint tanulási stratégiáinak vizsgálata. A következő kutatási kérdésekre kerestük a választ:

1. Hogyan változik a résztvevők aktivitása a kurzus során?
2. Az egyes változók milyen mértékben jelzik előre a kurzus záróvizsgáján nyújtott teljesítményt?
3. Milyen tanulói profilok azonosíthatók a tananyagban eltöltött idő, a tananyagban való előrehaladás és az elért teljesítmény alapján?

### **Módszerek**

#### *A kurzus áttekintése*

A kutatás keretében vizsgált MOOC-kurzus (Újratervezés Program 3.0) során a résztvevők alapvető digitális készségeket sajátíthattak el, illetve egy weboldal megalkotásán keresztül betekintést kaphattak a programozás alapjaiba is. Az online program elvégzésére a regisztrációt követően négy hét állt a rendelkezésükre. A négyhetes időtartam lehetőséget nyújtott arra, hogy a kurzus munka mellett, akár csak hétvé-

genként is elvégezhető legyen. A kurzus 14 modulból épült fel, az egyes modulok videós tananyagokból, valamint az ezek alapján megoldandó tesztek, feladatokból álltak. Az egyes modulok logikailag egymásra épültek, ugyanakkor egy adott modul elvégzése nem volt feltétele egy következő modul elvégzésének. Ennek következtében, ha egy résztvevő előzetesen már rendelkezett informatikai ismeretekkel, lehetősége volt bizonyos tananyagrészek átugrására. A kurzus záróvizsgával zárult, amelynek sikeres teljesítése után a résztvevőknek lehetőségük nyílt egy támogatott, államilag elismert végzettséget nyújtó IT-képzésen való részvételre.

A kurzus 2022. május-június hónapokban valósult meg. A kurzusra 9984 fő regisztrált, de a kurzus megvalósítására szolgáló NEXIUS Learning keretrendszerbe már csak 8479 fő jelentkezett be, és közülük 7760 fő kezdte meg a tananyagok feldolgozását. Az elemzésekbe kizárólag ez utóbbi résztvevők adatai kerültek be.

A résztvevők személyes adatait a felnőttképzésről szóló 2013. évi LXXVII. törvény előírásai alapján kezeltük, és kizárólag azokat az adatokat kértük be tőlük, amelyekre a törvény felhatalmazást ad. A kurzusra történő jelentkezéskor a résztvevők beleegyeztek abba, hogy anonimizált adataik kutatási célból felhasználhatók. Az elemzésekbe ezen kívül azokat a kontextuális adatokat vontuk be, amelyekre az oktatási keretrendszer lehetőséget nyújtott. Az elemzésekbe bevont változók nevét és leírását az 1. táblázat tartalmazza.

<i>Változók neve</i>	<i>Változók leírása</i>
tananyaggal eltöltött idő	a résztvevők által a tananyag tanulásával és a feladatok megoldásával eltöltött összes idő percben
tananyagban való előrehaladás	a résztvevők által elvégzett tananyag és az összes tananyag aránya
életkor	a résztvevők életkora években
iskolai végzettség	a résztvevők legmagasabb iskolai végzettsége: ISCED 1: befejezett általános iskolai 4. osztály, ISCED 2: befejezett általános iskolai végzettség, ISCED 3: középfokú végzettség, ISCED 4: technikum, ISCED 5: felsőoktatási szakképzés, ISCED 6-8: felsőfokú végzettség
angol nyelvtudás	a résztvevők nyelvtudása önbevallás alapján: 1: nincs, 2: alapfokú, 3: középfokú, 4: felsőfokú
teljesítmény	az elért pontszám és az összes pontszám aránya

1. táblázat: Az elemzésekbe bevont változók neve és leírása (saját szerkesztés)

A kiugró értékek azonosítására (Tukey 1977) vezette be a boxplot módszert, amelyhez a kvartiliseket alkalmazta. Az interkvartilis tartomány (*interquartile range, IQR*) az alsó ( $Q1$ ) és a felső ( $Q3$ ) kvartilis közötti különbség. A szabály a megfigyeléseket kiugrónak nyilvánítja, ha azok a

$$(Q1-k \cdot IQR; Q3+k \cdot IQR) \tag{1}$$

intervallumon kívül esnek, ahol  $k = 1,5$ . A boxplot nagyon népszerű módszer a kiugró értékek azonosítására (Shein és Fitrianto 2017), olyannyira, hogy az egyik legismertebb statisztikai szoftver, az SPSS is ezt használja (Meyers, Gamst és Guarino 2013). Kutatásunkban a Hoaglin és Iglewicz (1987) által alkalmazott  $k = 2,2$  értéket használtuk a küszöb meghatározásához, amely pontosabb eredményt szolgáltat (Banerjee és Iglewicz 2007). A változók közül a tananyaggal eltöltött idő esetén detektáltuk 30 résztvevő (0,39%) kiugró értékeit. A kiugró értékeket winsorizálással alakítottunk át, amely Charles P. Winsor biostatistikusról kapta a nevét. A winsorizálás lényege, hogy a kiugró értékeket a legnagyobb, illetve legkisebb még nem kiugró értékkel helyettesítjük (Wilcox 2017). Például a

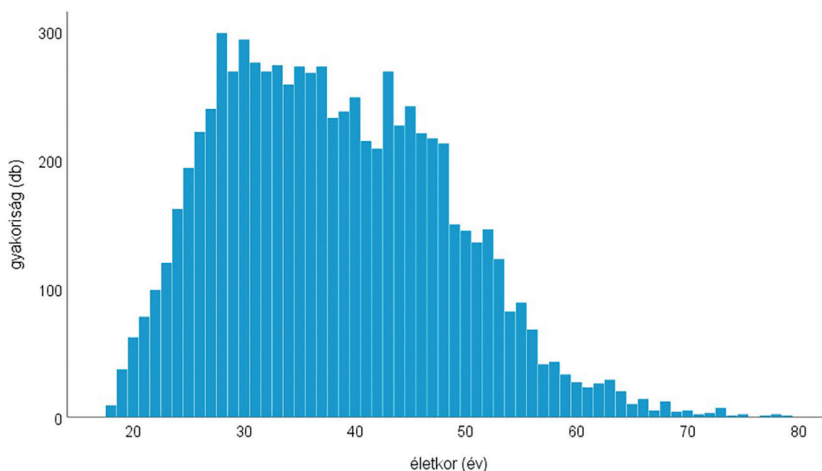
$$3, 3, 4, 4, 4, 4, 5, 5, 6, 14 \tag{2}$$

értékek esetén a 14 kiugró érték, emiatt az azt megelőző legnagyobb, még nem kiugró értékkel, azaz 6-tal helyettesítjük, így a

$$3, 3, 4, 4, 4, 4, 5, 5, 6, 6 \tag{3}$$

értékeket kapjuk. A winsorizálás egy olyan robusztus módszer a kiugró értékek kezelésére, amely jól használható a társadalomtudományokban (Blaine 2018).

### Demográfiai sokszínűség



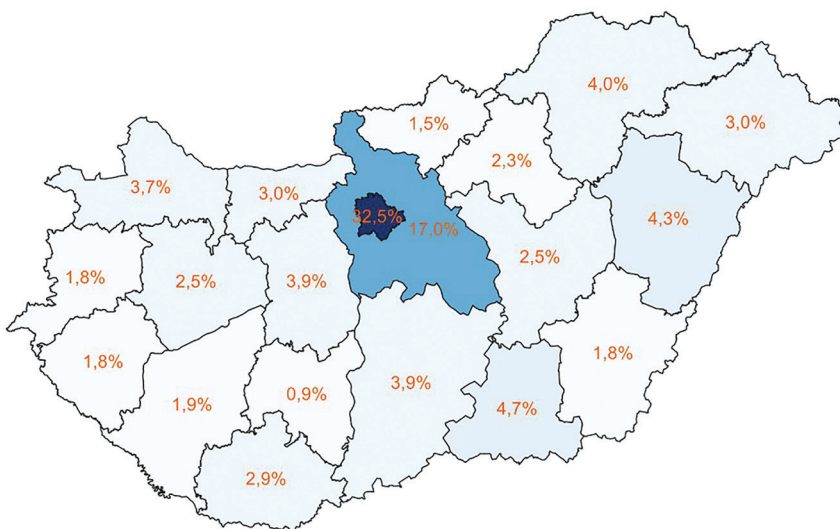
1. ábra: A résztvevők életkor szerinti eloszlása (saját szerkesztés)



A kurzusra jelentkezők közül a legfiatalabb résztvevő 18, míg a legidősebb 79 éves volt, az átlagéletkor  $M = 38,1$  ( $SD = 10,2$ ) év. Az életkori eloszlásnak két maximuma figyelhető meg, egyrészt 30 év, másrészt 45 év környékén (1. ábra). A maximumok egybeesnek Magyarország korfáján a két legnépesebb munkaképes korú korosztállyal (KSH 2023a).

A résztvevők többségének a legmagasabb iskolai végzettsége középfokú vagy felsőfokú volt. A résztvevők között nem volt ISCED 1 végzettségű, 1,6%-uk általános iskolai végzettségű (ISCED 2) volt, 35,7% középfokú végzettségű (ISCED 3), 8,9% technikumot (ISCED 4) végzett. 16,3%-uk legmagasabb iskolai végzettsége felsőoktatási szakképzés (ISCED 5) volt, míg 37,6%-nak volt felsőfokú végzettsége (ISCED 6-8).

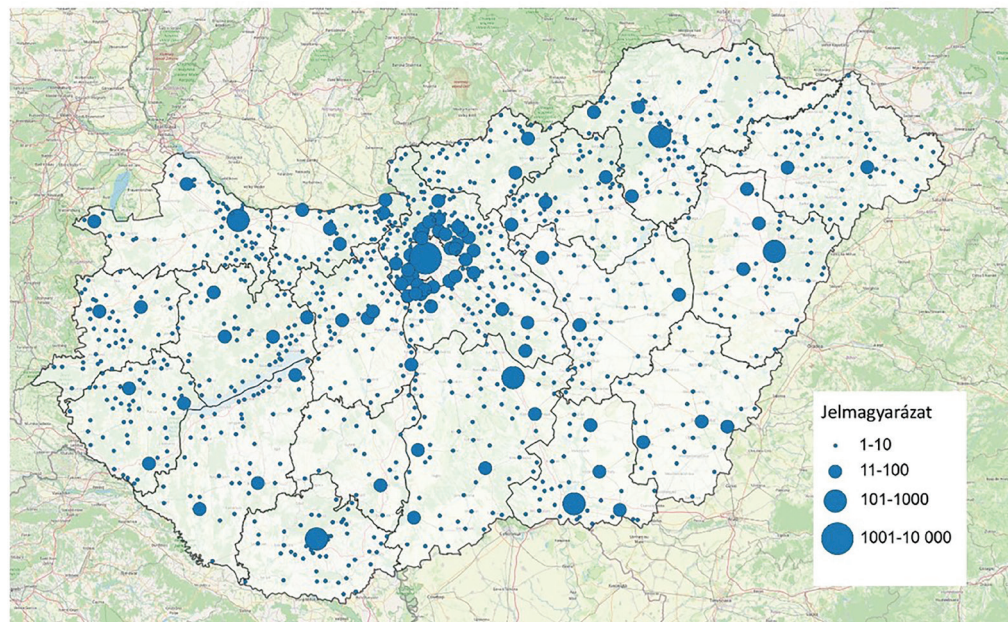
A résztvevők nyelvtudását önbevallás útján mértük, döntő többségük alapfokú (37,2%) vagy középfokú (42,9%) angol nyelvtudással rendelkezett. Az angol nyelvtudás nélküliek (6,3%) és a felsőfokú nyelvtudással rendelkezők (13,6%) jelentősen kisebb arányban voltak. Az angol nyelvtudásnak a kurzus szempontjából azért van jelentősége, mivel az informatika nyelve az angol, így feltételezésünk szerint a biztos angol nyelvismerettel rendelkezők könnyebben értik meg a tananyagot.



2. ábra: A résztvevők megyei szintű eloszlása (a százalékos értékek a résztvevők megyei arányát jelölik; saját szerkesztés)

A résztvevők földrajzi eloszlását lakóhelyük, illetve tartózkodási helyük alapján elemeztük. Amelyik résztvevő kizárólag lakóhelyet adott meg, annál azt az adatot vettük figyelembe, aki lakóhelyet és tartózkodási helyet is, annál az utóbbi került bele az elemzésekbe. A résztvevők közel fele (49,5%) Budapesten vagy Pest megyében lakik (2. ábra), ami lényegesen nagyobb a két megyének az országhoz viszonyított 31,3%-os lakosságárányánál (KSH 2023b). A kurzus résztvevői 1039 településen élnek (3. ábra). A legtöbb résztvevő budapesti volt (2519 fő), majd létszámban néhány megyeszékhely és Budapest agglomerációs települései következtek. Közülük

a legnagyobb résztvevői létszám is egy nagyságrenddel kisebb volt a budapestinél (Szeged, 216 fő). Résztvevői létszám tekintetében a másik végletet azon települések adják, ahonnan csak egy résztvevő volt (471 település). Ezek között számos kistelepülés van, amelyek sokszor csak néhány száz fős lakossággal rendelkeznek. Ezen a ponton is visszaigazolódott a MOOC-ok egyik nagy előnye, hogy akár a legtávolabbi, legkisebb településekről is elérhetőek a képzések (Námesztovszki et al. 2016).



3. ábra: A résztvevők települési szintű eloszlása (saját szerkesztés)

## Eljárások

A résztvevők kurzus során nyújtott aktivitásának vizsgálata érdekében az egyes modulok teljesítésének arányát vizsgáltuk. A kurzus záróvizsgáján elért teljesítmény előrejelzése érdekében többváltozós regresszióanalízist hajtottunk végre. Ennek során a tananyaggal eltöltött idő, a tananyagban való előrehaladás, az életkor, az iskolai végzettség és az angol nyelvtudás mint prediktív változók szerepét vizsgáltuk.

A résztvevők csoportjaink meghatározása érdekében látens profilelemzést végeztünk (Collins és Lanza 2010), amely egy modellalapú statisztikai módszer. A látens profil elemzés elvégzése során a résztvevők olyan csoportjait azonosítottuk, akik hasonló mintázatot mutattak a kurzuson való részvétel során. Az elemzésbe a tananyagban eltöltött időt, a tananyagban való előrehaladást, és a kurzus záróvizsgáján elért teljesítményt vontuk be.

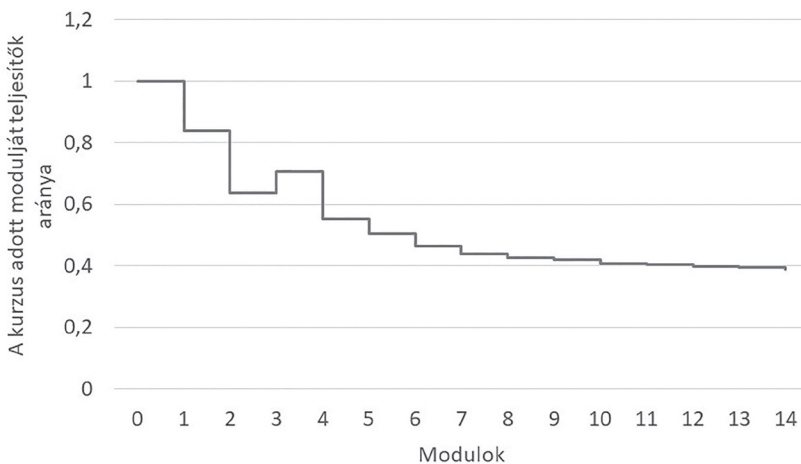
A hagyományos klaszterezési módszerekhez hasonlóan a látens profil elemzés során is kulcsfontosságú a látens csoportok optimális számának meghatározása. Erre többféle kritérium is alkalmazható, (1) a relatív illeszkedésindexek, (2) az entrópia

és (3) a Lo-Mendell–Rubin valószínűségi hányados teszt. A relatív illeszkedésindexek közé tartozik az AIC (*Akaike Information Criterion*, Akaike Információs Kritérium), a BIC (*Bayesian Information Criterion*, Bayesi Információs Kritérium) és az aBIC (*adjusted Bayesian Information Criterion*, korrigált Bayesi Információs Kritérium). A relatív illeszkedésindexek értékének csökkenésével növekszik a modell illeszkedése (Dziak et al. 2012). Az entrópia a csoportok homogenitásáról ad információt, azaz azt mutatja meg, hogy az adott modell mennyire pontosan sorolja az egyes résztvevőket az egyes profilokba. Az entrópia értéke 0 és 1 közötti értéket vehet fel, a 0,8 fölötti értékek megfelelő pontosságot jeleznek (Weller, Bowen és Faubert 2020). A Lo–Mendell–Rubin valószínűségi hányados teszt (*Lo-Mendell-Rubin Adjusted Likelihood Ratio Test*) arra használható, hogy összehasonlítsuk az  $n$  számú és az  $n-1$  számú látens osztályt tartalmazó modellt. A szignifikáns  $p$  érték azt jelzi, hogy az aktuálisan tesztelt  $n$  számú látens osztályt tartalmazó modell jobb illeszkedésű, mint az  $n-1$  számú látens osztályt tartalmazó modell. Az elemzéseket az Mplus 8.4 szoftverrel végeztük.

## Eredmények

*Az 1. kutatási kérdéshez kapcsolódó eredmény: Hogyan változik a résztvevők aktivitása a kurzus során?*

A kutatás során megvizsgáltuk, hogy mekkora azon résztvevők aránya, akik a kurzus adott modulját teljesítették (4. ábra). Breslow és munkatársai (2013) túlélési függvényként definiálták az így kapott görbét, mivel a függvény azt mutatja meg, hogy a résztvevők mekkora arányban vannak még jelen a kurzusban.



4. ábra: Az előrehaladás arányának változása az egyes modulok során (saját szerkesztés)

(Jelmagyarázat: a modulok esetén a 0 a kurzus kezdő időpontját, az 1-14 pedig az egyes modulok végét jelöli)

A 2. és a 3. modul közötti növekedéstől eltekintve a túlélési függvény folyamatosan csökkenő tendenciát mutat. A kezdeti 100%-ról az 1. modul végére 84%-ra csökken. A csökkenés a kurzus első felében erőteljes, több esetben 15-20 százalékpont körüli az egyes modulok között, a függvény értéke a kurzus felénél 44%. A kurzus második felében a csökkenés azonban lelassul, nem haladja meg a 1,5 százalékpontot az egyes modulok között, és a 14. modul után 39% körüli értéket vesz fel.

*A 2. kutatási kérdéshez kapcsolódó eredmény: Az egyes változók milyen mértékben jelzik előre a kurzus záróvizsgáján elért teljesítményt?*

A kutatási kérdés megválaszolása érdekében többváltozós regresszióanalízist végeztünk a vizsgált változók bevonásával, amelynek eredményei a 2. táblázatban láthatók.

<i>Független változók</i>	<i>r</i>	<i>β</i>	<i>r·β·100</i>	<i>p</i>
tananyaggal eltöltött idő	0,76	0,04	2,72	< 0,001
tananyagban való előrehaladás	0,94	0,91	84,90	< 0,001
életkor	-0,05	-0,01	0,05	0,017
iskolai végzettség	0,05	-0,01	-0,03	0,175
angol nyelvtudás	0,05	0,02	0,10	< 0,001
Teljes megmagyarázott variancia	87,74			

*2. táblázat: A kurzus záróvizsgáján elért teljesítmény mint függő változóra számított regresszióanalízis eredménye (saját szerkesztés)*  
*Megjegyzés: F = 11 081,965; p < 0,001*

A 2. táblázat a bevont független változók egyenkénti és együttes megmagyarázott varianciáját mutatja. A független változók a kurzus záróvizsgáján elért teljesítmény varianciájának közel 88%-át magyarázzák ( $R^2 = 87,74\%$   $p < 0,001$ ). Legnagyobb mértékben a tananyagban történő előrehaladás jelzi előre a kurzus záróvizsgáján elért teljesítményt. Ehhez képest nagyságrenddel kisebb mértékben bír előrejelző erővel a tananyagban töltött idő. Elhanyagolható az életkor és az angol nyelvtudás szerepe, míg az iskolai végzettség szerepe nem szignifikáns.

*A 3. kutatási kérdéshez kapcsolódó eredmény: Milyen tanulói proflok azonosíthatók a tananyagban eltöltött idő, a tananyagban való előrehaladás és az elért teljesítmény alapján?*

A látens profil elemzések illeszkedési mutatóit a 3. táblázat tartalmazza. Az AIC, a BIC és az aBIC is folyamatos csökkenést mutat a látens osztályok számának növekedése során. Az entrópia a két és három látens osztály esetén mutatott maximumot, de ezután is jelentősen meghaladta a megfelelőnek tekintett 0,8-es értéket. Az

LMR teszt (Lo–Mendell–Rubin valószínűségi hányados teszt) alapján a kilenc látens profillal rendelkező modellt fogadtuk el. A bevont változók alapján egyértelműen azonosíthatók a különböző típusú résztvevők. A 4. táblázat az adott látens csoporthoz tartozás valószínűségét mutatja.

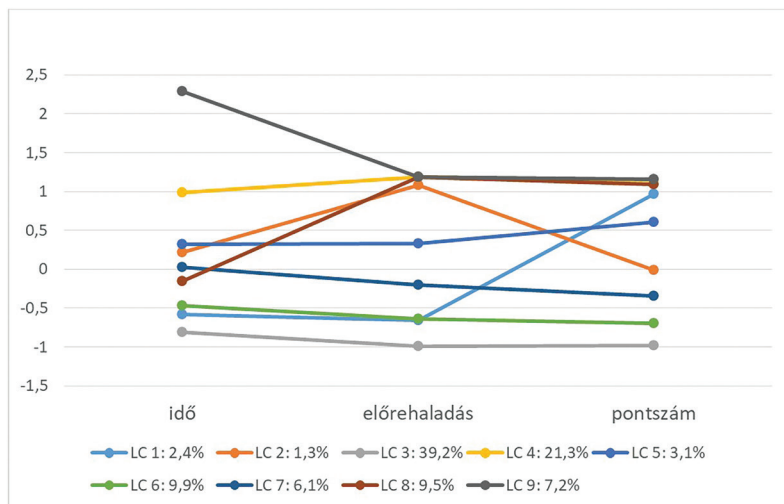
<i>Látens osztályok száma</i>	<i>AIC</i>	<i>BIC</i>	<i>aBIC</i>	<i>Entrópia</i>	<i>LMR teszt</i>	<i>p</i>
2	35 328	35 398	35 366	0,99	29 920	< 0,000
3	28 598	28 695	28 651	0,98	6555	< 0,000
4	25 930	26 055	25 998	0,94	2603	< 0,000
5	22 485	22 638	22 568	0,98	3359	< 0,000
6	19 702	19 883	19 800	0,95	2614	< 0,000
7	17 984	18 193	18 098	0,95	1679	< 0,000
8	16 241	16 477	16 369	0,94	1704	< 0,000
9	14 965	15 229	15 108	0,94	1524	< 0,000
10	13 896	14 188	14 055	0,94	1047	0,554

3. táblázat: A látens profil elemzések illeszkedési mutatói (saját szerkesztés)

<i>Látens csoport</i>	<i>Látens csoport</i>								
	<i>LC1</i>	<i>LC2</i>	<i>LC3</i>	<i>LC4</i>	<i>LC5</i>	<i>LC6</i>	<i>LC7</i>	<i>LC8</i>	<i>LC9</i>
LC1	0,995	0,000	0,000	0,000	0,002	0,003	0,001	0,000	0,000
LC2	0,000	0,938	0,000	0,020	0,003	0,000	0,000	0,039	0,000
LC3	0,000	0,000	0,985	0,000	0,000	0,015	0,000	0,000	0,000
LC4	0,000	0,002	0,000	0,939	0,001	0,000	0,000	0,036	0,022
LC5	0,002	0,005	0,000	0,006	0,975	0,000	0,012	0,000	0,001
LC6	0,000	0,000	0,083	0,000	0,000	0,896	0,021	0,000	0,000
LC7	0,001	0,000	0,000	0,000	0,010	0,032	0,957	0,000	0,000
LC8	0,000	0,003	0,000	0,097	0,001	0,000	0,000	0,899	0,000
LC9	0,000	0,000	0,000	0,095	0,000	0,000	0,000	0,000	0,905

4. táblázat: Az adott látens profillal jellemezhető csoporthoz tartozás valószínűsége (saját szerkesztés)

Az egyes látens csoportok jellemzését a nagy számuk miatt nem számozásuk szerinti, hanem logikai sorrendben tesszük meg. Az 5. ábra a kilenc, egymástól különböző látens csoport (LC) jellemzőit, az 5. táblázat az egyes csoportokhoz tartozó átlagos időt, előrehaladást és pontszámot mutatja.



5. ábra: A látens csoportok alakulása (saját szerkesztés)

(Jelmagyarázat: LC1: alacsony aktivitás magas teljesítménnyel; LC2: közepes idő, magas előrehaladás, közepes pontszám; LC3: legalacsonyabb aktivitás arányos teljesítménnyel;

LC4: legmagasabb aktivitás arányos teljesítménnyel; LC5: magas aktivitás arányos teljesítménnyel; LC6: alacsony aktivitás arányos teljesítménnyel; LC7: közepes aktivitás arányos teljesítménnyel; LC8: közepes idő, magas előrehaladás, magas pontszám; LC9: nagyon sok idő, magas előrehaladás, magas pontszám)

Látens csoport	idő (perc)		előrehaladás (%)		pontszám (%)	
	M	SD	M	SD	M	SD
LC1	421,0	327,2	18,4	10,1	84,2	12,0
LC2	1412,6	839,0	94,4	7,5	42,9	16,3
LC3	140,9	159,9	4,0	3,8	2,3	4,2
LC4	2306,4	455,1	99,3	2,7	92,0	9,7
LC5	1496,3	791,2	61,6	9,9	68,9	19,8
LC6	559,5	321,7	19,7	4,6	14,7	6,7
LC7	1141,7	565,0	38,3	6,9	29,2	9,2
LC8	890,9	430,7	99,2	3,0	89,5	11,2
LC9	3906,0	470,2	99,3	3,0	92,2	10,3

5. táblázat: Az egyes látens csoportokhoz tartozó átlagos idő (perc), előrehaladás (%) és pontszám (%) (saját szerkesztés)

---

A látens csoportokat két nagyobb kategóriába sorolhatjuk aszerint, hogy vizsgateljesítményük arányos volt-e a befektetett erőfeszítéssel (tananyaggal töltött idő és előrehaladás). Az LC3, LC4, LC5, LC6 és LC7 csoportoknál a vizsgateljesítmény arányos a befektetett erőfeszítéssel. A résztvevők közel négyötöde ezekbe a látens csoportokba tartozik (79,6%). Az ide tartozó résztvevőkre az jellemző, hogy akik több időt töltöttek a tananyaggal és nagyobb mértékű volt az előrehaladásuk benne, azok jobb eredményt értek el. Az ezeket a csoportokat jellemző görbék egymással nagyjából párhuzamosan, közel vízszintesen futnak. A résztvevők több mint harmada (3041 fő, 39,2%) teljes mértékben elvesztette a motivációját a képzéssel kapcsolatban (LC3). Ők minimális időt töltöttek a tananyaggal, minimális volt az előrehaladásuk, és vizsgát sem tettek, vagy ha igen, akkor igen alacsony teljesítményt mutattak. A résztvevők több mint ötöde (1652 fő, 21,3%) stabilan motivált maradt a képzés során (LC4). Ők viszonylag sok időt töltöttek a tananyaggal, amin nagyrészt végighaladtak, és jó eredményt értek el a vizsgán. A többi csoport létszáma az eddigiékhöz képest jóval kisebb arányú, minden esetben 10% alatt marad. Három látens csoport, az ötödik (LC5: 242 fő, 3,1%), a hatodik (LC6: 765 fő, 9,9%) és a hetedik (LC7: 475 fő, 6,1%) az előzőek között helyezkedik el a tananyaggal töltött idő, az előrehaladás és az elért pontszám tekintetében is.

Az LC1, LC2, LC8 és LC9 csoportoknál a vizsgateljesítmény nem volt arányos a tananyaggal töltött idővel és/vagy a tananyagban való előrehaladással. Az ezeket a csoportokat jellemző görbék egymással nem párhuzamosak, és nem is vízszintesen futnak. Az első csoportba (LC1: 184 fő, 2,4%) tartozó résztvevők nagyon rövid időt töltöttek a tananyaggal és nagyon alacsony volt az abban való előrehaladásuk, ennek ellenére nagyon jó eredményt értek el. A második csoportba tartozók (LC2: 102 fő, 1,3%) közepes idő alatt nagyrészt végighaladtak a tananyagon, ennek ellenére közepes vizsgaeredményt értek el. A nyolcadik csoportba tartozók (LC8: 739 fő, 9,5%) viszonylag rövid idő alatt nagyrészt végighaladtak a tananyagon, és nagyon magas pontszámot értek el a vizsgán. A kilencedik csoportba tartozók (LC9: 560 fő, 7,2%) extrém sok idő alatt haladtak végig a tananyagon és így értek el jó eredményt.

## Diszkusszió

A vizsgált MOOC-ot 7760 fő kezdte el, sikeres vizsgát pedig 2143 résztvevő tett, ami 72,4%-os lemorzsolódási arányt jelent. A lemorzsolódási arány lényegesen alacsonyabbnak bizonyult, mint a szakirodalmi adatok alapján nemzetközi szinten tapasztalható 90% körüli érték (Gregori et al. 2018; Hew és Cheung 2014; Morris 2013). Feltételezésünk szerint ennek egyik oka, hogy a sikeres vizsgázóknak lehetőségük volt a 100%-os támogatással megvalósuló, államilag elismert végzettséget adó informatikai képzésekre pályázni.

A kurzusban egy adott modul elvégzése nem volt feltétele egy következő modul elvégzésének, emiatt előfordulhatott, hogy egy résztvevő kihagyott egy modult, majd a következőt elvégezte. Ilyen eset látható a 2. és a 3. modul között, sok résztvevő hagyta ki a 2. modult, majd közülük többen elvégezték a harmadikat, így az utóbbi esetben magasabb a modul elvégzők aránya. A túlélési függvényt esetünkben azzal

az eltéréssel kell alkalmazni, hogy egy adott modul be nem fejezése után a résztvevő még továbbléphetett a következő modulra, tehát nem feltétlenül jelentkezik lemorzsolódóként, aki nem fejezett be egy adott modult. A résztvevők aktivitása a 2. és a 3. modul közötti növekedéstől eltekintve folyamatosan csökkenő tendenciát mutatott. A csökkenés a kurzus első felében volt nagyobb mértékű, az egyes modulok között több esetben 15-20 százalékpont körül volt. A kurzus második felében a csökkenés azonban lelassult, az egyes modulok között nem haladta meg a 1,5 százalékpontot. Azok a résztvevők, akik eljutottak a kurzus feléig, közel 90%-os valószínűséggel aktívak maradtak a kurzus végéig. Ez alapján a kurzus első felében nagyobb annak a valószínűsége, hogy a résztvevők nem végeznek el egyes modulokat, és ezáltal lemorzsolódnak a képzésről, azaz a kurzus első felében lényeges MOOC elem a résztvevők személyes megszólítása, motiválása, amivel várhatóan csökkenteni lehet a lemorzsolódási arányt. Az eredmények összhangban vannak korábbi kutatási eredményekkel. Breslow és munkatársai (2013) szintén a kurzus első felében detektáltak jelentősebb lemorzsolódást. Singer (2019) közel 300 000 HarvardX-re regisztrálót kérdezett meg kérdőíves módszerrel, hogy mi volt a céljuk a kurzusra való jelentkezés során. A megkérdezettek közel harmada nem válaszolt, majd az ő lemorzsolódási görbéjük csökkent a legdrasztikusabban. Hipotézise szerint ők azok, akik valójában nem is szándékozták elvégezni a kurzust.

A tananyaggal eltöltött idő, a tananyagban való előrehaladás, az életkor, az iskolai végzettség és az angol nyelvtudás záróvizsgán elért teljesítményre gyakorolt hatásának vizsgálatára többváltozós regresszióanalízist alkalmaztunk. A legnagyobb mértékben a tananyagban történő előrehaladás (84,9%) bírt előrejelző erővel a vizsgaeredményre. A tananyagban eltöltött idő ehhez képest egy nagyságrenddel kisebb előrejelző erővel bírt (2,72%). Az életkor (0,05%) és az angol nyelvtudás (0,10%) szerepe pedig még alacsonyabb hatással rendelkezett, míg az iskolai végzettség nem játszott szignifikáns szerepet az eredmény előrejelzésében. Az eredmények alapján megállapítható, hogy a résztvevők demográfiai adatai (életkor, iskolai végzettség, angol nyelvtudás) gyakorlatilag nincsenek hatással a kurzuson elért eredményességükre, azt szinte kizárólag a tanulási stratégiájuk határozza meg.

Chiu és munkatársai (2018) taiwani MOOC-kurzusok elemzése során arra a következtetésre jutott, hogy a végleges osztályzat a megválaszolt kérdések számával korrelált leginkább. Hasonló eredményre jutottak Moreno-Marcos és munkatársai (2020), akik két MOOC adatainak elemzése alapján megállapították, hogy a teljesítmény legjobb prediktorai a feladatokkal összefüggő változók (például megoldott feladatok, helyes feladatok aránya). Az elemzés során külön az egyes feladatok szerepét nem tudtuk vizsgálni, mert erre nem nyújtott lehetőséget a keretrendszer, viszont a feladatok megoldása beletartozott az előrehaladásba, így eredményeink összhangban vannak a szakirodalmi adatokkal.

A kontextuális adatokon alapuló látens profil elemzések lehetővé tették a résztvevők kurzuson mutatott aktivitásának elemzését. A szakirodalmi elemzés alapján azt feltételeztük, hogy a résztvevőknek legalább három csoportja azonosítható a logadatok alapján. A látens profil elemzés ennél jóval részletesebb eredményt adott, a résztvevők kilenc csoportját azonosította. A kilenc csoport két nagyobb kategóriába sorolható. Az egyik kategóriába tartozó csoportok résztvevői esetében a vizsgán



---

nyújtott teljesítmény arányos volt a befektetett erőfeszítéssel, azaz a tananyaggal töltött idővel és az előrehaladással. A másik kategóriába tartozó csoportok résztvevői esetén ez az arányosság nem állt fenn.

Az első kategóriába sorolható látens csoportokba (LC3, LC4, LC5, LC6, LC7) tartozott a résztvevők döntő többsége (79,6%). Az ezekben a csoportokban tartozó résztvevőkre jellemző, hogy ha több időt töltöttek a tananyaggal és jobban előrehaladtak benne, akkor jobb eredményt értek el a záróvizsgán. A 3. látens csoportba (LC3: legalacsonyabb aktivitás arányos teljesítménnyel) tartozó résztvevők (39,2%) teljes mértékben elvesztették a kurzus iránti érdeklődésüket, a legkevesebb idő alatt ( $M = 140,9$  perc,  $SD = 159,9$ ) a legkevésbé haladtak előre a tananyagban ( $M = 4,0\%$ ,  $SD = 3,8$ ), és ennek megfelelően a legrosszabb eredményt érték el ( $M = 2,3\%$ ,  $SD = 4,2$ ). A 4. látens csoportba (LC4: legmagasabb aktivitás arányos teljesítménnyel) tartozott a résztvevők 21,3%-a, ők ebben a kategóriában a leghosszabb idő alatt ( $M = 2306,4$  perc,  $SD = 455,1$ ) a legtovább jutottak a tananyagban ( $M = 99,3\%$ ,  $SD = 2,7$ ), és ennek megfelelően a legjobb eredményt érték el ( $M = 92,0\%$ ,  $SD = 9,7$ ). Aktivitás és teljesítmény szempontjából is közöttük helyezkedik el a 6. látens csoport (LC6: 9,9%, alacsony aktivitás arányos teljesítménnyel), a 7. látens csoport (LC7: 6,1%, közepes aktivitás arányos teljesítménnyel), és az 5. látens csoport (LC5: 3,1%, magas aktivitás arányos teljesítménnyel).

A második kategóriába sorolható látens csoportokba (LC1, LC2, LC8, LC9) a résztvevők ötöde tartozott (20,4%). Az ezekben a csoportokban tartozó résztvevőkre az jellemző, hogy aktivitásuk nem volt arányos a záróvizsgán nyújtott eredményükkel. Az 1. látens csoportba tartozó résztvevők (LC1: 2,4%) nagyon alacsony aktivitással jellemezhetők, azaz nagyon kevés időt töltöttek a tananyagban ( $M = 421,0$  perc,  $SD = 327,2$ ), és nagyon keveset is haladtak előre benne ( $M = 18,4\%$ ,  $SD = 10,1$ ), ennek ellenére jó eredményt értek el a záróvizsgán ( $M = 84,2\%$ ,  $SD = 12,0$ ). Az ebbe a csoportba tartozó résztvevőknek az iskolai végzettsége nem volt magasabb a többi csoport résztvevőinél. Az elért eredmény azzal magyarázható, hogy ők már rendelkeztek a tananyaggal kapcsolatos előzetes ismeretekkel, így anélkül is sikeresek tudtak lenni a vizsgán, hogy végighaladtak volna a tananyagban. A 2. látens csoportba tartozó résztvevők (LC2: 1,3%) közepes idő alatt ( $M = 1412,6$  perc,  $SD = 839,0$ ) szinte teljesen végigérték a tananyagban ( $M = 94,4\%$ ,  $SD = 7,5$ ), és közepes eredményt értek el ( $M = 42,9\%$ ,  $SD = 16,3$ ). Az ebbe a csoportba tartozó résztvevők esetén nem a tananyagban való előrehaladás, hanem a tananyaggal eltöltött idő határozta meg az eredményességet, amelynek vélhetően a figyelmetlen tanulás a magyarázata. A 8. látens csoportba tartozó résztvevők (LC8: 9,5%) a közepesnél kicsit rövidebb idő alatt ( $M = 890,9$  perc,  $SD = 430,7$ ) haladtak végig a tananyagban ( $M = 99,2\%$ ,  $SD = 3,0$ ), viszont ők jó eredményt értek el ( $M = 89,5\%$ ,  $SD = 11,2$ ). Ennek lehetséges magyarázata, hogy az ebbe a csoportba tartozó résztvevők jó képességekkel rendelkeznek, vagy előzetes ismereteik voltak a tananyaggal kapcsolatban. A 9. látens csoportba tartozó résztvevők (LC9: 7,2%) nagyon sok időt töltöttek el a tananyaggal ( $M = 3906,0$  perc,  $SD = 470,2$ ), miközben végighaladtak rajta ( $M = 99,3\%$ ,  $SD = 3,0$ ), és ezzel jó eredményt értek el ( $M = 92,2\%$ ,  $SD = 10,3$ ). A viselkedésük lehetséges magyarázata, hogy az ebbe a csoportba tartozó résztvevők kevésbé jó képességűek, emiatt töltöttek több időt a feladatok megoldásával. Egy másik lehetséges magyarázat az, hogy számukra

nagyon fontos volt a vizsgán való jó teljesítmény, és a sok tanulással biztosra akartak menni.

## Limitáció

A kutatás legfőbb korlátjának az tekinthető, hogy a kurzusnak helyet biztosító keretrendszer viszonylag korlátozott számú kontextuális változót rögzített. A kutatásunk során a tananyagban eltöltött időt és az abban való előrehaladást vizsgáltuk, míg a tanulmányok nagy része ennél több kontextuális változót elemez. A kutatás másik korlátja, hogy egy olyan kurzust vizsgáltunk, amely alapozó képzésként szolgált egy hagyományos rendszerű képzéshez, a kurzuson jól teljesítők közül kerülhettek ki azok, akik támogatott, bizonyítványt nyújtó informatikai képzésen vehettek részt. A támogatott, emiatt a résztvevő számára térítésmentes, államilag elismert képzés lehetősége vélhetően növelte a kurzust sikeresen teljesítők arányát, valamint befolyásolhatta a sikerességet. Ezen limitációk miatt a kutatás eredményei nem általánosíthatók az összes hazai és nemzetközi MOOC-kurzusra.

## Konklúzió

A tanulmányunk megerősíti számos korábbi kutatás eredményét. A résztvevők aktivitása, kurzus iránti elkötelezettsége a MOOC-kurzus során csökkent. Az aktivitás csökkenése a kurzus első felében volt jelentős, a második felében már csak minimálisan csökkent tovább. A résztvevők megtartására emiatt a kurzus első felében kell nagyobb figyelmet fordítani. Ekkor érdemes célzott üzenetekkel megkeresni a nem megfelelő aktivitású résztvevőket, hogy csökkenteni lehessen a kurzusról való lemorzsolódást. A kurzus záróvizsgáján való teljesítményt leginkább a tananyagban való előrehaladás határozza meg.

A kontextuális adatokon alapuló látens profil elemzés kilenc tanulói profilt azonosított. Ezek két nagyobb kategóriába sorolhatók, az egyik esetén a vizsgán elért teljesítmény arányos volt a kurzus során mutatott aktivitással, a másik esetén nem. A látens profilok alapján azonosíthatók a résztvevők hasonló tulajdonságokkal rendelkező csoportjai, amelyek megalapozhatják a személyre szabott oktatás lehetőségét.

## Irodalom

Banerjee, Sharmila és Boris Iglewicz. "A Simple Univariate Outlier Identification Procedure Designed for Large Samples." *Communications in Statistics: Simulation and Computation* 36, no. 2 (2007): 249–263.

<https://doi.org/10.1080/03610910601161264>

- 
- Barthakur, Abhinava, Vitomir Kovanovic, Srecko Joksimovic, George Siemens, Michael Richey és Shane Dawson. "Assessing Program-Level Learning Strategies in MOOCs." *Computers in Human Behavior* 117, (2021): 106674.  
<https://doi.org/10.1016/j.chb.2020.106674>
- Blaine, Bruce E. "Winsorizing." In Bruce B. Frey (Szerkesztő). *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*, 1817–1818. Thousand Oaks, California : SAGE Publications, Inc., 2018.
- Breslow, Lori, David Pritchard, Jennifer DeBoer, Glenda Stump, Andrew Ho és Daniel Seaton. "Studying Learning in the Worldwide Classroom: Research into EdX's First MOOC." *Research in Practice and Assessment* (2013)
- Chiu, Yu-Chen, Hwai-Jung Hsu, Jungpin Wu és Don-Lin Yang. "Predicting Student Performance in MOOCs Using Learning Activity Data." *Journal of Information Science and Engineering* 34 (2018): 1223–1235.  
[https://doi.org/10.6688/JISE.201809\\_34\(5\).0007](https://doi.org/10.6688/JISE.201809_34(5).0007)
- Class central. "A Decade of MOOCs: A Review of MOOC Stats and Trends in 2021." Utolsó hozzáférés: 2023. október 2.  
<https://www.classcentral.com/report/moocs-stats-and-trends-2021/>
- Collins, Linda M. és Stephanie T. Lanza. *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*. Wiley Series in Probability and Statistics. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2010.  
<https://doi.org/10.1002/9780470567333>
- Dziak, John J., Donna L. Coffman, Stephanie T. Lanza és Runze Li. "Sensitivity and Specificity of Information Criteria." *Technical Report Series* no. 12-119, (2012).
- Fan, Yizhou, Jelena Jovanović, John Saint, Yuhang Jiang, Qiong Wang és Dragan Gašević. "Revealing the Regulation of Learning Strategies of MOOC Retakers: A Learning Analytic Study." *Computers and Education* 178, (2022)  
<https://doi.org/10.1016/j.compedu.2021.104404>
- Gregori, Pablo, Vicente Martínez és Julio José Moyano-Fernández. "Basic Actions to Reduce Dropout Rates in Distance Learning." *Evaluation and Program Planning* 66, (2018): 48–52.  
<https://doi.org/10.1016/j.evalprogplan.2017.10.004>
- Henning, Christian, Marina Meila, Fionn Murtagh és Roberto Rocci. *Handbook of Cluster Analysis*. New York: Chapman and Hall/CRC, 2015.  
<https://doi.org/10.1201/b19706>
- Hew, Khe Foon és Wing Sum Cheung. "Students' and Instructors' Use of Massive Open Online Courses (MOOCs): Motivations and Challenges." *Educational Research Review* 12, (2014): 45–58.  
<https://doi.org/10.1016/j.edurev.2014.05.001>
- Hoaglin, David C. és Boris Iglewicz. "Fine-Tuning Some Resistant Rules for Outlier Labeling." *Journal of the American Statistical Association* 82, no. 400 (1987): 1147–1149.  
<https://doi.org/10.2307/2289392>
- Ho, Andrew, Justin Reich, Sergiy Nesterko, Daniel Seaton, Tommy Mullaney, Jim Waldo és Isaac Chuang. "HarvardX and MITx: The First Year of Open Online Courses." *HarvardX and MITx Working Paper* no. 1 (2014).

- Hollands, Fiona és Devayani Tirthali. *MOOCs: Expectations and Reality*. New York: Center for Benefit-Cost Studies of Education, Teachers College, Columbia University, 2014.
- Kassymova, G., A. Bekalaeva, D. Yershmanova, N. Flindt, T. Gadirova és Duisenbayeva Sh. “E-Learning Environments and Their Connection to the Human Brain.” *International Journal of Advanced Science and Technology* 29, no. 9 (2020): 947–954.
- Khalil, Mohammad és Martin Ebner. “Clustering Patterns of Engagement in Massive Open Online Courses (MOOCs): The Use of Learning Analytics to Reveal Student Categories.” *Journal of Computing in Higher Education* 29, no. 1 (2017): 114–132.  
<https://doi.org/10.1007/s12528-016-9126-9>
- Központi Statisztikai Hivatal. “Magyarország népességének száma nemek és életkor szerint, január 1.” Utolsó hozzáférés: 2023. október 2. a  
<https://www.ksh.hu/interaktiv/korfak/orszag.html>
- Központi Statisztikai Hivatal. “A lakónépesség nem, vármegye és régió szerint, január 1.” Utolsó hozzáférés: 2023. október 2. b  
[https://www.ksh.hu/stadat\\_files/nep/hu/nep0034.html](https://www.ksh.hu/stadat_files/nep/hu/nep0034.html)
- Longstaff, Emily. “How MOOCs Can Empower Learners: A Comparison of Provider Goals and User Experiences.” *Journal of Further and Higher Education* 41, no. 3 (2017): 314–327.  
<https://doi.org/10.1080/0309877X.2015.1100715>
- Majó-Petri Zoltán, Prónay Szabolcs, Huszár Sándor és Dinya László. “Digitális transzformáció az egyetemeken – Egy tömeges, nyílt, online oktatási működési modell, és az egyetemisták digitális oktatáshoz fűződő attitűdjének vizsgálata.” *Információs Társadalom* XX, 1. szám (2020): 72–94.  
<https://doi.org/10.22503/INFTARS.XX.2020.1.4>
- Matcha, Wannisa, Dragan Gašević, Jelena Jovanović, Nora’Ayu Ahmad Uzir, Chris W. Oliver, Andrew Murray és Danijela Gasevic. “Analytics of Learning Strategies: The Association with the Personality Traits.” In *ACM International Conference Proceeding Series*, 151–160. Association for Computing Machinery, New York, United States: LAK ’20: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, 2020  
<https://doi.org/10.1145/3375462.3375534>
- Meyers, Lawrence S., Glenn C. Gamst és A. J. Guarino. *Performing Data Analysis Using IBM SPSS*. Hoboken, New Jersey: John Wiley & Sons, 2013.
- Molnár Gyöngyvér. “Az IKT Szepe a Felsőoktatás Megújításában.” *Magyar Tudomány* 182, 11. szám (2021): 1488–1501.
- Molnár Gyöngyvér, Turcsányi-Szabó Márta és Kárpáti Andrea. “Digitális forradalom az oktatásban – perspektívák és dilemmák.” *Magyar Tudomány* 181, 1. szám (2020): 56–67.  
<https://doi.org/10.1556/2065.181.2020.1.6>
- Moreno-Marcos, Pedro, Ting-Chuen Pong, Pedro Merino és Carlos Delgado-Kloos. “Analysis of the Factors Influencing Learners’ Performance Prediction With Learning Analytics.” *IEEE Access* 8, (2020): 5264–5282.  
<https://doi.org/10.1109/ACCESS.2019.2963503>
- Morris, Libby V. “MOOCs, Emerging Technologies, and Quality.” *Innovative Higher Education* 38, no. 4 (2013): 251–252.  
<https://doi.org/10.1007/s10755-013-9263-2>

- 
- Muthén, Linda K. és Bengt O. Muthén. *Mplus User's Guide*. 7th ed. Los Angeles, CA: Muthén & Muthén, 2012.
- Námesztovszki Zsolt, Balázs P. Dorotya, Kovács Cintia, Major Lenke és Dijana Katuovic. "Tanulói Aktivitás Mintázatai Három MOOC Képzés Alapján." *Információs Társadalom* XVI, 4. szám (2016): 40-60.  
<https://doi.org/10.22503/inftars.XVI.2016.4.3>
- Námesztovszki Zsolt, Esztelecki Péter, Kőrösi Gábor, Major Lenke és Dragana Glušac. "Tapasztalatok Három Saját Készítésű MOOC Kapcsán." *Információs Társadalom* XV, 3. szám (2015): 63-84.  
<https://doi.org/10.22503/inftars.XV.2015.3.4>
- Ortega-Arranz, Alejandro, Erkan Er, Alejandra Martínez-Monés, Miguel L. Bote-Lorenzo, Juan I. Asensio-Pérez és Juan A. Muñoz-Cristóbal. "Understanding Student Behavior and Perceptions toward Earning Badges in a Gamified MOOC." *Universal Access in the Information Society* 18, no. 3 (2019): 533-549.  
<https://doi.org/10.1007/s10209-019-00677-8>
- Rodrigues, Rodrigo Lins, Jorge L. C. Ramos, Joao C. Sedraz Silva, Alex Sandro Gomes, Fernando Da Fonseca De Souza és Alexandre Magno Andrade Maciel. "Discovering Level of Participation in MOOCs through Clusters Analysis." In *2016 IEEE 16th International Conference on Advanced Learning Technologies (ICALT)*, 232-233. Austin, TX, USA: International Conference on Advanced Learning Technologies (ICALT), 2016.  
<https://doi.org/10.1109/ICALT.2016.45>
- Romero, Cristobal és Sebastian Ventura. "Educational Data Mining and Learning Analytics: An Updated Survey." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* (2020): 1-21.  
<https://doi.org/10.1002/widm.1355>
- Shein, Hui Wong és Anwar Fitrianto. "A Comparative Study of Outliers Identification Methods in Univariate Data Set A Comparative Study of Outliers Identification Methods in Univariate Data Set." *Advanced Science Letters* 23, (2017): 1422-1427.  
<https://doi.org/10.1166/asl.2017.8366>
- Singer, Judith D. "Reshaping the Arc of Quantitative Educational Research: It's time to Broaden Our Paradigm." *Journal of Research on Educational Effectiveness* 12, no. 4 (2019): 570-593.  
<https://doi.org/10.1080/19345747.2019.1658835>
- Szani Ferenc és Merkovity Norbert. "Javaslat egy közép-kelet európai képzési piac kialakítására." *Információs Társadalom* XIV, 1. szám (2014): 91-108.  
<https://doi.org/10.22503/inftars.XIV.2014.5>
- Tóth, Krisztina, Heiko Rölke, Frank Goldhammer és Ingo Barkow. "Educational Process Mining: New Possibilities for Understanding Students' Problem-Solving Skills." In Benő Csapó és Joachim Funke (Szerkesztők). *The Nature of Problem Solving. Using Research to Inspire 21st Century Learning*, 193-209. Paris: OECD, 2017.  
<https://doi.org/10.1201/9781003160618-1>
- Tukey, John W. *Exploratory Data Analysis*. London: Addison-Wesley, 1977.
- Valle, Rodrigo Del és Thomas M. Duffy. "Online Learning: Learner Characteristics and Their Approaches to Managing Learning." *Instructional Science* 37, no. 2 (2009): 129-149.  
<https://doi.org/10.1007/s11251-007-9039-0>

- Weller, Bridget E., Natasha K. Bowen és Sarah J. Faubert. “Latent Class Analysis: A Guide to Best Practice.” *Journal of Black Psychology* 46, no. 4 (2020): 287–311.  
<https://doi.org/10.1177/0095798420930932>
- Wilcox, Rand R. “Winsorized Robust Measures.” In N. Balakrishnan, T. Colton, B. Everitt, W. Piegorisch, F. Ruggeri és J.L. Teugels (Szerkesztők). *Wiley StatsRef: Statistics Reference Online*, 2017.  
<https://doi.org/10.1002/9781118445112.stat06339.pub2>
- Williams, Kyle M, Rose E. Stafford, Stephanie B. Corliss és Erin D. Reilly. “Examining student characteristics, goals, and engagement in Massive Open Online Courses.” *Computers & Education* 126, (2018): 433–442.  
<https://doi.org/10.1016/j.compedu.2018.08.014>
- Yu, Chen-Hsiang, Jungpin Wu és An-Chi Liu. “Predicting Learning Outcomes with MOOC Clickstreams.” *Education Sciences* 9, no. 2 (2019)  
<https://doi.org/10.3390/educsci9020104>
- Zhu, Meina, Annisa R. Sari és Mimi Miyoung Lee. “Trends and Issues in MOOC Learning Analytics Empirical Research: A Systematic Literature Review (2011–2021).” *Education and Information Technologies* 27, no. 7 (2022).  
<https://doi.org/10.1007/s10639-022-11031-6>