



Közzététel: 2024. február 12.

A tanulmány címe:

Most vegyem, vagy később? A repülőjegyek árazásával kapcsolatos összefüggések elemzése

Szerző:

MARTUS BETTINA

a Budapesti Corvinus Egyetem Adatelemzés és Informatika Intézete Statisztika Tanszékének egyetemi adjunktusa

E-mail: bettina.martus@uni-corvinus.hu

DOI: <https://doi.org/10.20311/stat2024.02.hu0187>

Az alábbi feltételek érvényesek minden, a Központi Statisztikai Hivatal (a továbbiakban: KSH) *Statisztikai Szemle* c. folyóiratában (a továbbiakban: Folyóirat) megjelenő tanulmányra. Felhasználó a tanulmány vagy annak részei felhasználásával egyidejűleg tudomásul veszi a jelen dokumentumban foglalt felhasználási feltételeket, és azokat magára nézve kötelezőnek fogadja el. Tudomásul veszi, hogy a jelen feltételek megszegéséből eredő valamennyi kárért felelősséggel tartozik.

1. A jogszabályi tartalom kivételével a tanulmányok a szerzői jogról szóló 1999. évi LXXVI. törvény (Szjt.) szerint szerzői műnek minősülnek. A szerzői jog jogosultja a KSH.
2. A KSH földrajzi és időbeli korlátozás nélküli, nem kizárólagos, nem átadható, térítésmentes felhasználási jogot biztosít a Felhasználó részére a tanulmány vonatkozásában.
3. A felhasználási jog keretében a Felhasználó jogosult a tanulmány:
 - a) oktatási és kutatási célú felhasználására (nyilvánosságra hozatalára és továbbítására a 4. pontban foglalt kivétellel) a Folyóirat és a szerző(k) feltüntetésével;
 - b) tartalmáról összefoglaló készítésére az írott és az elektronikus médiában a Folyóirat és a szerző(k) feltüntetésével;
 - c) részletének idézésére – az átvevő mű jellege és célja által indokolt terjedelemben és az eredetihez híven – a forrás, valamint az ott megjelölt szerző(k) megnevezésével.
4. A Felhasználó nem jogosult a tanulmány továbbértékesítésére, haszonszerzési célú felhasználására. Ez a korlátozás nem érinti a tanulmány felhasználásával előállított, de az Szjt. szerint önálló szerzői műnek minősülő mű ilyen célú felhasználását.
5. A tanulmány átdolgozása, újra publikálása tilos.
6. A 3. a)–c) pontban foglaltak alapján a Folyóiratot és a szerző(ke)t az alábbiak szerint kell feltüntetni:
„*Forrás: Statisztikai Szemle* c. folyóirat 102. évfolyam 2. számában megjelent, **Martus Bettina** által írt, **Most vegyem, vagy később? A repülőjegyek árazásával kapcsolatos összefüggések elemzése** című tanulmány (link csatolása)”
7. A Folyóiratban megjelenő tanulmányok kutatói véleményeket tükröznek, amelyek nem feltétlenül esnek egybe a KSH vagy a szerzők által képviselt intézmények hivatalos álláspontjával.

Martus Bettina

Most vegyem, vagy később? A repülőjegyek árazásával kapcsolatos összefüggések elemzése

Should I buy now or later? An analysis of the correlations of airplane ticket pricing

Martus Bettina, a Budapesti Corvinus Egyetem Adatelemzés és Informatika Intézete Statisztika Tanszékének egyetemi adjunktusa
E-mail: bettina.martus@uni-corvinus.hu

A repülőjegyek ára számtalan tényező hat: például az ún. rejtett városok megléte, a versenytársak aktivitása és számossága, egy gazdasági, társadalmi vagy környezeti kihívás, de akár az is, hogy a nyári időszakban, vagy azon kívül vesszük meg a jegyet. A fogyasztók, illetve a repülőtársaságok modelljeit el kell különíteni egymástól, mert míg az egyik réteget a repülőjegy minél alacsonyabb áron való beszerzése motiválja, addig a másik a legnagyobb nyereség elérését célozza meg. Kutatásomban friss adatok alapján azt elemeztem, hogy mely tényezők hatottak leginkább a repülőjegyárakra a Budapest–New York útvonal esetében. Eredményeim szerint jelentősen csökkentette a jegyárát, ha ugyanaz a repülőtársaság szolgáltatja az oda-vissza utat, valamint az is hatással volt a jegyárakra, hogy melyik időszakban vásárolták a repülőjegyet.

Kulcsszavak: repülőjegy, árazás, interakció

Air ticket prices are affected by several factors: the presence of so-called hidden cities phenomena, the activity and number of competitors, an economic, social, environmental challenge, or even just the fact that we buy the ticket during the summer period or before. The consumer's and airline's models must be separated from each other, because while one segment is motivated by the purchase of the plane ticket at the lowest possible price, the other aims at achieving the most profit. We analyzed based on recent data which factors had the greatest effect on the prices of plane tickets by examining the Budapest - New York route. Based on our results, we found that the matching of airlines on the round trip significantly reduced the prices of plane tickets, and the period/time of purchase of the plane ticket also had an effect on the ticket price.

Keywords: airplane ticket, pricing, interaction

Minden nyár előtt aktuálissá válik a kérdés: ha repülős utazást tervezünk, mikor érdemes megvenni a repülőjegyet, hogy a lehető legolcsóbban jussunk hozzá? Ez az a probléma, amelyre a repülőutakra készülők nagyon szeretnék tudni a megoldást, ugyanakkor nincs rá jó válasz. Aki utazott már repülőgéppel és kellőképpen árérzékeny, tudja, hogy nagyon sok tényező játszik szerepet a repülőjegyárak alakulásában.

Néhány tényező a fogyasztók számára is nyilvánvaló, de valójában számos változó határozza meg az árak alakulását, köztük egészen meglepőek is. Például egyes feltevések szerint minél többször nézünk meg egy adott repülőjegyet az interneten, végül annál drágábban kapjuk meg. *Aslani, Modarres és Sibdari (2014)* szintén feltették a mindenkit foglalkoztató kérdéseket: Az oda-vissza útra szóló repülőjegy miért olcsóbb, mint az egy irányba, egy útra szóló? Hogy lehet az, hogy egy hosszabb út megtétele kevesebbe kerül, ellenben ha csak az átszállási helyig repülünk, a jegyárat tekintve drágábban jönnénk ki?

Az árazás eleve nehéz feladat, ha pedig ezzel kapcsolatos előrejelzést szeretnénk adni, az még bonyolultabb dolog. Egyrészt az árazás komplex, több stratégiát is magában foglaló tevékenység, másrészt az adatok gyakran titkosak (egyes légitársaságok árazási modellje nem nyilvános), és több terület megközelítési módját kell ötvözni, így az értékesítést, a marketingét, a pénzügyét és a pszichológiáját. A cégek számára fontos, hogy a költségeik megtérüljenek, hogy az adott repülőjegyeket eladják a vevőknek, marketingstratégiával pozicionálják magukat, és ösztönözzék, meggyőzzék a vásárlókat, hogy az adott pillanatban a legjobb áron veszik meg a repülőjegyet.

Az előrejelzés nemcsak azért bonyolult, mert nagyon sok tényezőt kell figyelembe venni, hanem azért is, mert az algoritmusok meghatározására olyan komplex módszerek is használhatók, mint például a *machine learning*, azaz a gépi tanulás, amely az adatok alapján azonosítja a mintákat, és ezeket felhasználva modelleket alkothatunk. Fel kell készülnünk azonban arra is, hogy ezzel sem tudunk mindent belekalkulálni a modellbe. A 2020-ban fellépő Covid19-világjárvány idején a *Világgazdaság (2020)* cikke szerint példátlan promóciókkal igyekeztek életben maradni a légitársaságok. Ez érthető is, mivel 2020-ban a 2019-es évhez képest a ki- és beutazók aránya 70%-kal csökkent a járvány terjedésének megfékezését célzó korlátozások és szigorú szabályok következtében. A becslések szerint a pandémia előtti, 2019. évi forgalmi szintet a légitársaságok csak 2024-ben lesz-

nek képesek újra elérni. A járvány idején azonban nem voltak szokatlanok a különféle repülőjegy-akciók, a „korlátlan bérletek”, illetve „a két szék egy áráért” promóciók (*Világgazdaság*, 2020).

2022-ben a világ kezdett visszatérni a normális kerékvágásba, jól látszott ez a tendencia az utasszámokon és a repülőjegyárakon is. 2023 februárjára a *Financial Times* (2023) elemzése szerint az átlagos repülőjegyár a világ több mint 600 legnépszerűbb úti célja esetében 27,4%-kal növekedett az egy évvel korábbihoz képest.¹ A KSH adatait vizsgálva (*KSH Stadat*, 2023) látható, hogy Magyarországon 2002 és 2022 között az induló utasok száma 2019-ig nőtt, az induló járatok számához hasonlóan, majd a Covid19-járvány hatása itt is érzékelhető volt, és drasztikus csökkenést hozott. Annak ellenére, hogy 2022-ben fellendülés következett be, sem az utasok, sem a járatok száma nem érte el a 2019. évi értékeket.

Tanulmányomban azt vizsgálom, hogy mely tényezők és hogyan befolyásolják a repülőjegyárak alakulását. Elemzésem során a Budapest–New York útvonalat vizsgáltam 2023. március 23. és augusztus 17. között. Fontos megjegyezni, hogy a vizsgálatnak nem célja előrejelzést adni az adatok alapján, csupán szeretném feltárni, hogy mely tényezők milyen hatással voltak az adott időszakban az árakra. Bár a nemzetközi szakirodalomban széles körben foglalkoznak ezzel a témával, Magyarországon kevésbé. Jelen tanulmányban egyrészt magyar viszonylatban vizsgálom a járatokat, másrészt áttekinthető és értelmezhető módon igyekszem bemutatni több mint 131 ezer megfigyelés alapján, hogy mely tényezők mennyiben járultak hozzá a repülőjegyárak alakulásához.

1. Szakirodalmi áttekintés

Mivel a repülőjegy-árazás dinamikus árazási stratégiát kíván, azaz az ár folyamatosan változik több tényező függvényében, érdemes csoportosítani a hozzá kapcsolódó tényezőket a következőképpen: költségalapú, versenyszerű, illetve keresletalapú árazás.

A 20. század folyamán a közlekedés szerepe felértékelődött, a légi közlekedés egyre több ember számára vált megfizethetővé, a logisztika pedig még nagyobb teret nyitott az egyes termékek, áruk elérhetőségében. Ahogy kinyílt a világ, a fajlagos költségek is változtak, mivel a cégek több vevőt is elérhetnek, a *költségálapúság* pedig a szolgáltatási szektor ártrendeződésével is szoros kapcsolatban van.

¹ Az elemzés a legnépszerűbb utakra, az átlagos egyirányú árakra vonatkozik, amelyekben az adók és egyéb díjak nincsenek benne.

A repülőjegyek árának elemzésekor érdemes tehát elsőként a szolgáltatások *költségeit* számba venni, megnézni a szolgáltatók szolgáltatásainak körét, ennek megfelelően pedig az általuk kínált szolgáltatás árát. *Dudás, Boros és Pál (2016)* egy 31 ezer repülőjegyet magában foglaló adatbázis adatai alapján több tényezőt vizsgált. Egyrészt azt, hogy volt-e olyan weboldal, amely az általuk vizsgált időszakban minden esetben olcsóbb jegyeket adott a többinél, másrészt a metakeresőket (Skyscanner, Kayak) és az online utazási irodákat (Expedia, Orbitz, Cheaptickets) hasonlították össze. Az eredmények alapján azt a következtetést vonták le, hogy nincs olyan weboldal, amely a többinél mindig olcsóbb repülőjegyeket tud biztosítani, ugyanakkor a metakeresők az online utazási irodákénál kedvezőbb jegyárakat kínálnak. Természetesen ennek oka a szolgáltatási körben keresendő: a metakeresők egy linkkel továbbírányítanak az adott weboldalra, ahol a jegyek megvehetők, míg az online utazási irodák oldalain eleve biztosított a jegyvásárlás lehetősége. *Law és szerzőtársai (2011)* a regionális és a helyi online utazási irodák által meghatározott repülőjegyeket vizsgálták a globális utazási irodákéhoz képest. Eredményeik alapján arra jutottak, hogy az általuk vizsgált időszakban és járatokra a globális szereplők árai magasabbak voltak, mint helyi társaiké.

A költségekhez kapcsolódó fontos tényező az árhoz kötődő percepció, amelyet *Yu (2008)* vizsgált. Kutatásom során online jegyárakat néztem, az internet naponta több száz ajánlatot is felhoz, ami aztán a fogyasztó percepcióit, véleményét is formálja. Fontos tehát hangsúlyozni, hogy ezek csupán percepciók, mégis lényeges szerepet tölthetnek be.

A keresletalapú árazásnál a fentieknél összetettebb helyzet alakul ki. A fogyasztói szokások megismerése lehetőséget ad arra, hogy minél jobban meghatározzuk a keresletet alakító tényezőket. Az egyik ilyen tényező az idő. *Dudás, Boros és Pál (2016)* arra jutottak, hogy az előszezonban a jegyek olcsóbbak, mint a főszezonban, és a repülőjegyek hétfőn és vasárnap magasabbak, mint a többi napon. *Dudás és munkatársai (2016)* a városok térkapcsolatainak komplexebbé válását és a termódosító hatásokat vizsgálták kvantitatív kutatási módszerekkel, a diszkont és a hagyományos légitársaságok eltérő üzletpolitikáját és gyakorlatát elemezve. Eredményeik rámutattak arra, hogy a repülőjegyek foglalási időpontjában az árak az idő függvényével U alakú görbén ábrázolhatók, valamint nem egyenletes csökkenő vagy növekvő tendencia figyelhető meg az árakban. *Law és szerzőtársai (2011)* a repülőjegyek napon belüli árváltozásait vizsgálták retúrjegyekre és rögzített indulásai dátummal öt online utazási irodán keresztül, ami alapján arra jutottak, hogy nem érdemes a jegyeket korán megvenni.

A nemzeti sajátosságok, vagy akár az egyes országokban az internethasználat jellemzői szintén olyan tényezők lehetnek, amelyek hatással vannak a repülőjegyeladásokra, mivel segítenek megismerni és alakítani a keresletet. *Ruiz-Mafe és*

szerzőtársai (2013) arra a kérdésre keresték a választ, hogy mely fő tényezők befolyásolják a repülőjegy-, azon belül is az online vásárlást. Tanulmányukban egy kérdőív segítségével vizsgálták az online vásárlókat, 25 holland és 25 spanyol vásárlót. Eredményeik azt mutatták, hogy a nemzeti kultúra hatással van az internetezési szokásokra, ezáltal az online repülőjegy-vásárlásokra is. A spanyoloknál az egyéni normák, az önkontroll és az attitűd pozitívan befolyásolja a repülőjegy-vásárlást, míg a hollandoknál az attitűd a fő tényező. Keresletről lévén szó, a motivációs és az érzelmi faktorok is jelentős szerepet töltenek be a kereslet tekintetében, amit *Abayi és Khoshtinat (2016)* vizsgált.

Alderighi, Gaggero és Piga (2022) bemutatták, hogy egy eszköz értékét általában három tényező határozza meg: a romlandóság, a szűkösség és az időközi árdiszkrimináció. A repülőjegyek esetében az ülések árait vizsgálták, mivel ezek, ha nem is kelnek el azonnal, be vannak árazva, és idővel eladják őket. Tanulmányukban azt írják, hogy a helyek árai az idő múlásával csökkennek, kivéve az indulás előtti néhány napos időszakot, illetve a repülőjegyek árai nőnek az egymás után eladott ülések áraival megegyezően, azaz minél több ülés lesz „eladva” a repülőn, annál drágább lesz a jegy. *Lewis (2021)* a kedvezményeket vette alapul a repülőjegyárak vizsgálatakor, azaz tágabb értelemben véve az árdiszkriminációt, valamint a verseny erősségét (*Lewis, 2021; Silva–Dias–Dos Reis, 2022; Kerkemeros et al., 2023; Fiore–Zanzalari, 2023*).

A keresletre bármilyen gazdasági, társadalmi, vagy akár környezeti jelenség hatással lehet, amely a repülőjegyek árazási stratégiáját is alaposan megváltoztathatja. *Morlotti és Redondi (2023)* a Covid19-járvány hatásait vizsgálták, eredményeik pedig azt igazolták, hogy a kereslet visszaesésének hatására minden főbb európai járatnál, amely Olaszországból indult, vagy Olaszországba tartott, 31%-kal csökkent az egy kilométerre jutó ár.

A visszatérő, nagy ügyfélkör az árazás szintjén a keresletre és a költségekre alapoz. *Koo, Mantin és O’Connor (2011)* azt vizsgálták, hogy a légitársaságok mennyire használják a saját honlapjukat jegyértékesítésre, illetve igénybe vesznek-e online utazási irodákat, ügynökségeket (*online travel agency, OTA*). Természetesen a jegyárakra ez a tényező is hatással van. Azok, akik nagy, megbízható vevőkörrrel rendelkeznek, kevésbé veszik igénybe az OTA-platformokat, ugyanis a kereslet várhatóan biztos, ugyanakkor pluszköltségek sem merülnek fel az OTA-k miatt.

Az árat nagymértékben meghatározza az egyes légitársaságok egymással való együttműködése is, erre jelentenek példát az olyan nagy szövetségek, mint a Star Alliance. A szövetségek révén lehetőség van arra, hogy a légitársaságok a költségeket csökkentsék, mégis a keresletnek megfelelő szolgáltatást tudjanak biztosítani. *Sampaio és Urdanoz (2022)* tanulmánya az amerikai piacot vizsgálta 2003 és 2019 között. Tanulmányukban rávilágítottak arra, hogy az együttműködések

száma az elmúlt időszakban drasztikusan visszaesett, valamint azt tapasztalták, hogy nem jelenthető ki egyértelműen, hogy az ugyanazon szövetségbe tartozó légitársaságok alacsonyabb árat fognak adni. Természetesen az együttműködés szintje és az út típusa szintén meghatározza az együttműködő és a nem együttműködő légitársaságok árait. Ugyanezt a témát *Ivaldi, Petrova és Urdanoz (2021)* ökonometriai módszerekkel vizsgálták az amerikai piacon. Eredményeik alapján látható, hogy átlagosan ott vannak alacsonyabb árak, illetve szórások az árakban, ahol a légitársaságok valamely szövetség tagjai.

A versenyszerű árazáshoz kapcsolódó tényező az ún. „rejtett város árazás”² (*hidden city pricing* vagy *hidden city ticketing*), vagyis amikor A-ból B-be közvetlenül eljutni drágább, mintha A-ból B-be egy köztes úti cél beiktatásával repülnénk (*Gaggero, Luttmann, 2023; Oh-Huh, 2022*). *Gaggero és Luttmann (2023)* 473 000 járatot vizsgáltak meg 2019. október 1. és 2019. december 31. között. Azt tapasztalták, hogy a „rejtett város árazás” akkor jelenik meg, amikor az adott út vonatkozásában verseny van a légitársaságok között, leginkább az utazás előtti héten, és a végállomás általában nem egy forgalmas csomópont. Jellemzően olyan légitársaságoknál fordul ez elő, amelyek több nagy csomópontot is lefednek (pl. United Airlines). Tekintettel arra, hogy online vesszük a repülőjegyet, előfordulhat, hogy keresés közben ugyanarra a járatra, ugyanazokkal a feltételekkel egy másik oldalon is találunk jegyet, de eltérő – akár kedvezőbb – árral.

A repülőjegy-árazási stratégiáknál alapvetően négy irányt kell megkülönböztetnünk, amelyekre *Abdella és szerzőtársai (2021)* is felhívják a figyelmet. A légitársaságok oldaláról kétfajta modell létezhet: a kereslet-előrejelzés és az árdiszkrimináció modellje, a fogyasztói oldalról pedig az optimális repülőjegy-vásárlási idő modellje, valamint a repülőjegyár-előrejelző modell. A két oldal érdeke eltér: a légitársaságok a bevételeiket akarják maximalizálni, a vevők pedig minél többet akarnak megtakarítani a repülőjegyük árán. Így minél inkább stratégiai szempontból közelítik meg a vevők a vásárlást, annál nehezebb lesz a légitársaságoknak a dinamikus árazási stratégiát alkalmazni és több profitot termelni. Ez viszont fordítva is igaz: minél inkább stratégiai szemléletű a légitársaság, annál nehezebb a dolga a vásárlónak. A két érdek között mindig van egy trade-off. Érdemes a modelleket kicsit jobban szemügyre venni. Számos tanulmány foglalkozott már azzal, hogy vagy a fogyasztók, vagy a cégek (légitársaságok) oldaláról vizsgálja meg ezeket a modelleket. De vajon milyen tényezők lesznek a magyarázóváltozói az egyik, illetve a másik modellnek, el lehet-e élesen különíteni ezeket, és milyen módszerekkel érdemes vizsgálni? Mivel az árazási stratégia a repülőjegyek ese-

² A „rejtett város árazás” egy olyan anomália, amikor egy köztes úti cél beiktatásával olcsóbban tudunk jegyet vásárolni, mintha a kiindulási helyről a célállomásig vennénk meg a közvetlen jegyet, illetve olyan esete is van, amikor a kiindulási hely és a célállomás közti repülőjegy drágább, mintha a célállomás köztes helynek minősülne (*Wang-Ye, 2012*).

tében nagyon összetett folyamat, érdemes egy minél komplexebb modellt felépíteni. Mivel azonban ezeknek a lehetősége korlátozott, egyrészt az adatok megszerzése, másrészt a rengeteg információ miatt, így érdemes az adatbányászathoz *machine learninget* alkalmazni. Ha jobban belegondolunk, számos tényező befolyásolja, illetve befolyásolhatja a repülőjegy árát. Egyfelől a légitársaságok közötti verseny már önmagában is meghatározó faktor, de az ünnepek, egy adott szezon vagy egy közelgő fontos esemény (pl. olimpiai játékok) mind-mind hatással lehet az áralakításra (Abdella et al., 2021; Martínez-Navarro-Trinquescoste, 2017). Koc és Arslan (2021) a *machine learning* és a hagyományos modellek alkalmazását egyaránt ajánlják. Tanulmányukban hét különböző modelltechnikát szimulálnak, négy különböző adatállományon.

Abdella és szerzőtársai (2021) szakirodalmi szemlét állítottak össze, amely az eddig elkészített árazási modelleket vizsgálta. Összegyűjtötték, hogy az egyes kutatások mely tényezőket vizsgálták a különböző modellekben (1. táblázat).

1. táblázat

A modellek és azok tényezőinek bemutatása*Presentation of the models and their factors*

Tényezők	Fogyasztók modelljei		Légitársaságok modelljei	
	optimális jegyvásárlási idő előrejelzése	jegyár-előrejelzés	kereslet-előrejelzés	ár-diszkrimináció
Járatok száma az adott napon	x			
Indulás előtti napok száma	x			x
Ajánlatkérés napja a héten belül	x			x
Jegy korábbi ára (minimum, maximum, átlag, Nash-egyensúlyi ár)	x	x	x	x
Légitársaság	x			
Út hossza	x			
Ajánlatok száma az adott napon	x			
Indulás helye	x			
Érkezés helye	x			
Indulás dátuma	x	x		
Visszaérkezés dátuma	x			
Indulás napja a héten	x		x	x
Indulás napja a hónapban	x			
Indulás napja az évben	x			
Kereslet (elmúlt időszak kereslete alapján)	x			x
Árak (ugyanazon korábbi időszakban, ugyanazon ütemezéssel)		x		

(A táblázat folytatása a következő oldalon)

(folytatás)

Tényezők	Fogyasztók modelljei		Légitársaságok modelljei	
	optimális jegyvásárlási idő előrejelzése	jegyár-előrejelzés	kereslet-előrejelzés	ár-diszkrimináció
Indulás városa		x		
Érkezés városa		x		
Jegy vásárlásának dátuma		x		
Légitársaság által üzemeltetett járatok száma			x	
Légitársaság teljesítménye (késések aránya, megállók száma, kapacitás)			x	
Lakosság jövedelme			x	
Fogyasztóiár-index			x	
Fogyasztói hívások száma (ügyfélszolgálat)			x	
Keresőmotor használata			x	
Előzetes foglalások száma				
Indulási idő napközben				
Jegy vásárlása a héten			x	x
Versenytárs akciója			x	
Fizetés módja				x
Jegy korlátozásai (vásárlási határidő, utazási korlátozások)			x	
Szabadnapok, ünnepnapok			x	
Rendkívüli események				x
Évszak hatása			x	

Forrás: *Abdella és munkatársai (2021)* alapján saját szerkesztés.

2. Felhasznált adatok, alkalmazott módszer

2.1. Az adatok bemutatása és a hipotézisek

Az adatok összegyűjtése a kiwi.com oldalról történt, ahonnan *webscraping* segítségével mindennap letöltöttem az adatokat egy Python-script megírásával. A scriptben beállítottam, hogy az adatok legyűjtése mindig este 8 órakor történjen. Ez azért volt fontos feltétel, mert ezzel elkerülhettem azt, hogy a különböző időpontokban való leszedés hatással legyen az árakra. *Abdella és munkatársai (2021)*

ugyanis tanulmányukban kifejtik, hogy a repülőjegyek esetében könnyen bekövetkezhet, hogy akár napi 7 alkalommal is megváltoznak az árak, tekintettel a dinamikus árazási stratégiára. A fenti beállítás viszont segít abban, hogy ezt a hatást kiszűrjem. További paraméter volt a kiindulási és az érkezési hely megadása. Csak a Budapest–New York útvonalon vizsgálódttam, a reptereket nem határoztam meg, így New Yorkban három repterre érkezhett a gép. A választás azért erre az útvonalra esett, mivel forgalmas repterekről van szó, és akár a magyar–amerikai gazdasági hatások is nyomon követhetők a repülőjegyek áraiban. Hosszú útról van szó, így az a kiinduló feltételezésem, hogy az árak is nagyobb mértékben volatilisak, a fogyasztók pedig nagyobb árérzékenységet mutatnak. Az árfolyamhatás kiszűrésére a jegyárakat amerikai dollárban kértem le. Mivel a közvetlen járatok jelentősége ezen a szakaszon általában csekély (csak egy légitársaság indít ilyen járatokat), az egyátszállásos járatokat néztem meg. Az indulási időnél hagytam egy kis játékteret, ugyanakkor kikötöttem, hogy a hálaadásra (2023. november 23.), amely fontos ünnepnek számít az USA-ban, „megérkezzünk”, így az indulás november 17. és 23. (23:59 percig) között bármikorra eshet. Retúr repülőjegyeket néztem, és ennek során nem a visszaindulás dátumát, hanem az ott töltött éjszakák számát adtam meg (7–14). További feltétel volt, hogy a teljes utazási idő egy irányba nem haladhatta meg a 17 órát, így több olyan légitársaság sem adott ajánlatot, amely egyébként közlekedik az adott útvonalon, csak nagyobb mértékű kitérével (például kelet felé indul, és ott van az átszállás, így túllépi a meghatározott teljes utazási időt). A lekérdezés definiálása során fontos volt azt a szempontot is megadni, hogy csak olyan jegyeket gyűjtsön le a rendszer, amelyek egy feladott poggyászt tartalmaznak, különben az árak emiatt is különbözhetek volna.

2023. március 23. és 2023. augusztus 17. között 131 455 megfigyelésem volt (2. táblázat). A szemléltetés kedvéért április minden napján megjelenítettem az ajánlatok számát, amiből jól látszik, hogy rengeteg ajánlatot kaphatunk az adott útvonalra az általunk megadott időszakra és feltételekkel, ugyanakkor voltak olyan napok – és ez minden hónapban előfordult –, hogy a vártnál kevesebb ajánlatot jelenített meg a rendszer (pl. április 2.). Fontos azt is hangsúlyozni, hogy bizonyos légitársaságok egyáltalán nem jelentek meg az ajánlatokban, ami fakadhat abból, hogy egyes járatok két átszállással közlekednek, illetve abból is, hogy aránytalanul hosszú időt jelentene az utazás, és ez a megadott feltételekbe nem fért bele. A megfigyelések alapján a leggyorsabb ajánlat 10 óra 55 perces utat kínált, míg a leghosszabb az általam meghatározott 17 óra volt.

2. táblázat

Az ajánlatok számának alakulása a vizsgált időszakban*The numbers of offers in the examined period*

Letöltés dátuma	Ajánlatok száma	Letöltés dátuma	Ajánlatok száma
Március	7 406	Április 18.	932
Április	25 582	Április 19.	502
Április 1.	781	Április 20.	837
Április 2.	362	Április 21.	846
Április 3.	929	Április 22.	855
Április 4.	925	Április 23.	942
Április 5.	883	Április 24.	960
Április 6.	819	Április 25.	871
Április 7.	915	Április 26.	880
Április 8.	944	Április 27.	853
Április 9.	821	Április 28.	906
Április 10.	823	Április 29.	865
Április 11.	833	Április 30.	832
Április 12.	793	Május	29 524
Április 13.	947	Június	23 529
Április 14.	948	Július	29 900
Április 15.	924	Augusztus	15 514
Április 16.	885	Összesen	131 455
Április 17.	969		

A vizsgálat célja annak feltárása, hogy a Budapest–New York retúr repülőjegy árára mely tényezők hatnak szignifikánsan, és milyen módon. Az értékeléshez többváltozós lineáris regressziós modellt használtam. Kutatásom során az alábbi kérdéseket vizsgáltam meg:

K1: Nyáron drágábban lehet-e jegyet vásárolni a Budapest–New York útvonalra, mint tavasszal?

K2: Ha végig ugyanazzal a légitársasággal repülünk, drágább-e a jegy (légitársaságtól függetlenül)?

K3: Az indulás előtt jóval korábban megvásárolt jegyek olcsóbbak-e, mint az indulás előtt rövid idővel vásároltak?

K4: Ha egy adott hónap elején vásároljuk meg a repülőjegyet, drágább jegyet fogunk-e kapni, mint a hónap közepén vagy végén?

K5: A hétvégén megvásárolt jegy drágább-e a hétköznap vásárolthoz képest?

K6: A kint töltött idő hossza hatással van-e a jegyárra?

K7: Az átszállás helyének tájolása befolyásolja-e a jegyárat?

K8: A korán, illetve a későn induló járatokra olcsóbbak-e a jegyek, mint a napközben indulókra?

K9: Az érkezés ideje hatással van-e a repülőjegyárakra, a korán vagy a későn érkező járatok jegyei drágábbak-e?

K10: A nagyobb lakosságú városokban történő átszállás miatt olcsóbbak-e a jegyek? (Itt a nagy lakosságszám feltételezi a repterek megnövekedett forgalmát és nagyságát is.)

K11: A rövidebb utazási idő drágább jegyet jelent-e?

K12: Ha az átszállás a kiutazás során más városban történik, mint a visszaút során, akkor a jegy olcsóbb?

K13: Ha különböző légitársaságokkal utazunk a teljes út során vagy egyes szakaszokon, olcsóbb-e a jegy (repülőtársaságtól függetlenül)?

K14: Az adott napra adott ajánlatok száma fordítottan arányos-e az árral, azaz, ha egy adott napon sok ajánlatot kapunk az adott feltételekre, akkor olcsóbbak-e a repülőjegyek?

A kérdések megválaszolását és a hipotézisek tesztelését az R statisztikai programban végeztem el, felhasználva az lmtree, a Sandwich, a Datarium és a Car programcsomagokat.

2.2. A modellváltozók definiálása

A bevezetésben bemutatott szakirodalmak az egyes modellekben már definiáltak endogén és exogén tényezőket, amelyek a repülőjegyek meghatározásánál szerepet játszottak. Mivel elemzésem arra a kérdésre irányul, hogy mely tényezők befolyásolják az adott feltételek alapján megkeresett repülőjegyek árát, eredményváltozóm az ár lesz, a magyarázóváltozókat pedig a 3. táblázat mutatja.

A hipotézisekhez kapcsolódóan 23 magyarázóváltozó került be a végső regressziós modellbe. Ezek nagyobb része kvalitatív, minőségi változó, így ezek dummyváltozóként szerepelnek a modellben. A táblázatban látható megjegyzések fontos korlátokra utalnak, mint például az árfolyam esetében. Az árfolyam a gazdasági hatások változásait igyekezett volna bemutatni a modellben, ugyanakkor mivel hétvégére nincs középárfolyam, mindenhol az adott heti pénteki árfolyamot használtam. Mivel minden hipotézisem az árra irányul, így ez a célváltozóm. A megfigyelések alapján azt az eredményt kaptam, hogy az adott időszakban a legalacsonyabb jegyár 617 amerikai dollár volt, míg a legmagasabb 6177,92. A hisztogramon kiválóan megfigyelhető (1. ábra), hogy egy jobbra elnyúló változóról beszélünk, amelyben outlier értékek is vannak.

3. táblázat

A magyarázóváltozók típusai és leírásuk
Types of the explanatory variables and their description

Változó neve	Típusa	Változó leírása	Mértékegysége	Megjegyzés
ajanlatok_szama	kvantitatív	Elérhető napi ajánlatok száma	darab	
lesz_honap	kvalitatív	Az ajánlat a hónap mely harmadában volt elérhető?	hónap eleje, közepe, vége	
lesz_hetk_hetv	kvalitatív	Az ajánlat hétköznap vagy hétvégén volt elérhető?	hétköznap, hétvége	
ejszaka	kvantitatív	New Yorkban töltött éjszakák száma	darab	
ind_nap_ossz_szama	kvantitatív	Indulás előtt hány nappal kértük le az ajánlatot?	nap	
napszak_1	kvalitatív	Az indulás ideje Budapestről, napon belül (napszak)	délelőtt, délután, hajnal	Nem volt este induló járat.
atsz_1_taj	kvalitatív	Budapesthez képest az átszállás helyének tájolása	kitérő, kitérő nélkül	A Magyarországhoz képest északnyugatról, nyugatról induló gépeket vettem kitérő nélkülieknek.
atsz_1_lak	kvantitatív	Budapestről indulva, az átszállás helyének lakosság száma	fő	
ind_1_hetk_hetv	kvalitatív	Budapestről az indulás hétköznap vagy hétvégén történik?	hétköznap, hétvége	
oda_utaz_ido_e	kvantitatív	Budapestről New Yorkba a teljes utazási idő	óra	
erk_2_napszak	kvalitatív	New Yorkba érkező járat érkezési napja	délelőtt, délután	
ind_3_hetk_hetv	kvalitatív	New Yorkból az indulás hétköznap vagy hétvégén történik?	hétköznap, hétvége	

(A táblázat folytatása a következő oldalon)

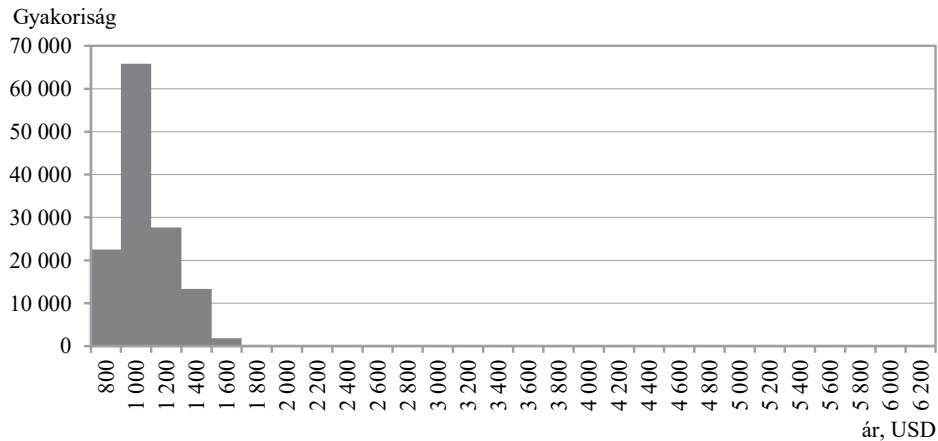
(folytatás)

Változó neve	Típusa	Változó leírása	Mértékegysége	Megjegyzés
ind_3_napszak	kvalitatív	Az indulás ideje New Yorkból, napon belül (napszak)	délután, este, hajnal	Nem volt délelőtt induló járat.
atsz_helye_oda_vissza	kvalitatív	Az átszállás helyének egyezése Budapestről és New Yorkból	egyeznek, nem egyeznek	
vissza_utazasi_ido_e	kvantitatív	New Yorkból Budapestre a teljes utazási idő	óra	
arfolyam	kvantitatív	A letöltés napján érvényes közép-árfolyam	HUF/USD	Hétfvégén középárfolyam nincs, így mindig a pénteki árfolyamot vettem alapul az adott hétfvégéhez.
atsz_legit_1_4	kvalitatív	Légitársaságok egyezősége az átszállások esetén	egyeznek, nem egyeznek	
atsz_legit_2_3	kvalitatív	Légitársaságok egyezősége a New Yorkba érkező és New Yorkból induló légitársaságok esetén	egyeznek, nem egyeznek	
atsz_legit_ossz	kvalitatív	Légitársaságok egyezősége minden repülés alkalmával	egyeznek, nem egyeznek	
nyar	kvalitatív	Az ajánlat nyár előtt vagy nyáron érkezett?	nyár előtt, nyáron	
ind_3_tajolas	kvalitatív	Budapesthez képest az átszállás helyének tájolása	kitérő, kitérő nélkül	A Magyarországhoz képest északnyugatról, nyugatról induló gépeket vettem kitérő nélkülieknek.
ind_3_lakosság	kvantitatív	New Yorkból indulva, az átszállás helyének lakosság száma	fő	
bp_erk_napszak	kvalitatív	Budapestre érkezés ideje (napszak)	délelőtt, délután, este	Nem volt hajnalban érkező járat.

Forrás: saját szerkesztés.

1. ábra

Az ár változó hisztogramja
The histogram of the price variable



2.3. Modellépítés

A függő és a független változók bemutatása után az alábbi kiinduló regressziós modellt határoztam meg:

$$\begin{aligned} \text{Ár} = & \beta_0 + \beta_1 \times \text{ajánlatok_szama} + \beta_2 \times \text{lesz_honap} + \beta_3 \times \text{lesz_hetk_hetv} + \beta_4 \times \text{ejszaka} \\ & + \beta_5 \times \text{ind_nap_ossz_szama} + \beta_6 \times \text{napszak_1} + \beta_7 \times \text{atsz_1_taj} + \beta_8 \times \text{atsz_1_lak} \\ & + \beta_9 \times \text{ind_1_hetk_hetv} + \beta_{10} \times \text{oda_utaz_ido_e} + \beta_{11} \times \text{erk_2_napszak} \\ & + \beta_{12} \times \text{ind_3_hetk_hetv} + \beta_{13} \times \text{ind_3_napszak} + \beta_{14} \times \text{atsz_helye_oda_vissza} \\ & + \beta_{15} \times \text{vissza_utazasi_ido_e} + \beta_{16} \times \text{arfolyam} + \beta_{17} \times \text{atsz_legit_1_4} + \beta_{18} \times \text{atsz_legit_2_3} \\ & + \beta_{19} \times \text{atsz_legit_ossz} + \beta_{20} \times \text{nyar} + \beta_{21} \times \text{ind_3_tájolás} + \beta_{22} \times \text{ind_3_lakosság} \\ & + \beta_{23} \times \text{bp_erk_napszak} + \varepsilon \end{aligned}$$

A fenti modell paramétereinek becslését a legkisebb négyzetek módszerével (*Ordinary Least Squares, OLS*) végeztem el. A modell használatának feltételeként, az autokorreláció vizsgálatokor érdemes tekintettel lenni arra, hogy vagy időbeli, vagy területi autokorrelációról beszélhetünk. A megfigyelések nem idősoros modellt alkotnak, hanem keresztmetszetit, mivel nem egy járat adatait, hanem többet figyelem meg ugyanazokban az időpontokban, több változóval, így a mintavételt függetlennek tekintjük, ahol ez a feltétel automatikusan megvalósul.

Az OLS modellfeltevései közül ki kell emelni még a homoszkedaszticitást (*Hunyadi, 2006*), amely megköveteli, hogy a hibatag feltételes varianciája legyen konstans, e nélkül a regressziós koefficiensek becslése nem határos, és a hiba kovarianciamátrixának becslése nem torzítatlan. A heteroszkedaszticitás akár fenn is

állhat, mivel ehhez akár az is elég lehet, hogy bizonyos időközönként új légitársaságok ajánlatot adjanak arra az útra, amelyet vizsgálunk. Nagy mintanagyságunk azonban lehetővé tette, hogy ezt a problémát robusztus kovarianciamátrix-bebecsléssel oldjuk meg. A p-értékek alapján a modellünk nem változott, minden magyarázóváltozó szignifikáns (5. táblázat).

A multikollinearitás vizsgálatához a varianciainfláló faktorokat vettem alapul (VIF-eket). A VIF-ek esetében a szakirodalomban nincs egységes küszöbszám az értelmezésre (Kovács, 2008). A VIF-ek reciprokai toleranciamutatókat jelölnek. Ezek minél közelebb vannak a nullához, annál nagyobb a multikollinearitás valószínűsége (Kovács–Petres–Tóth, 2004). A vizsgálatban nincs 7,64-nál nagyobb VIF-érték, ami bizonyos szakirodalmak alapján már erős, zavaró multikollinearitást jelentene, ugyanakkor mivel ez nem egységes érték, és a mintanagyság is megfelelő hozzá, valamint a toleranciamutató 0,1 felett marad, így egy magyarázóváltozó kivétele sem indokolt a modelltől.

4. táblázat

A magyarázóváltozókhoz kapcsolódó VIF-ek és toleranciamutatók

VIFs and tolerance indices related to the explanatory variables

Magyarázóváltozó	VIF	Toleranciamutató	Magyarázóváltozó	VIF	Toleranciamutató
ajanlatok_szama	1,120	0,893	ind_3_napszak	3,095	0,323
lesz_honap	1,274	0,785	atsz_helye_oda_vissza	1,788	0,559
lesz_hetk_hetv	1,005	0,995	vissza_utazasi_ido_e	1,622	0,617
ejszaka	1,020	0,981	arfolyam	1,141	0,877
ind_nap_ossz_szama	4,571	0,219	atsz_legit_1_4	5,401	0,185
napszak_1	1,391	0,719	atsz_legit_2_3	7,641	0,131
atsz_1_taj	1,831	0,546	atsz_legit_ossz	5,788	0,173
atsz_1_lak	1,840	0,543	nyar	4,506	0,222
ind_1_hetk_hetv	1,070	0,934	ind_3_tajolas	1,837	0,544
oda_utaz_ido_e	2,414	0,414	ind_3_lakosság	1,762	0,567
erk_2_napszak	1,840	0,544	bp_erk_napszak	3,456	0,289
ind_3_hetk_hetv	1,088	0,919			

Forrás: saját szerkesztés.

A regressziós modell felírásakor a sok magyarázóváltozó léte miatt létrehozható lett volna egy főkomponens-analízis is, ugyanakkor, ha a változóimat mesterséges változókká alakítom, akkor a modell átláthatósága és magyarázhatósága csökken. A magyarázóváltozók sokasága a megfigyelések számossága, nagysága miatt vizsgálható a modellben.

3. Eredmények

3.1. A többváltozós lineáris regressziós modell eredményei

Talán meglepő lehet, de a többszörös determinációs együttható értéke 0,38, azaz a magyarázóváltozók együttese 38%-ban magyarázzák az árak varianciáját. Ennek az aránynak számos oka lehet. Meg kell említeni, hogy előfordulnak például rendszerhibás árak (*error fares*), amikor a repülőjegyár irreálisan olcsó, mondjuk egy elírás miatt. Az is megeshet, hogy egy adott légitársaság marketingokokból úgy dönt, hogy akciót hirdet, így a repülőjegyárat ismét csak egy olyan tényező befolyásolja, amely a modellünkben nem szerepel, mégis hatással van rá. Egy természeti katasztrófa szintén hatással lehet a repülőjegyek árának alakulására, gondoljunk csak a 2023 nyarán bekövetkező görögországi tüzesetekre. Látható tehát, hogy a repülőjegy-árazási folyamat nagyon összetett, amelyre csekély mértékben lehet rálátásunk, ugyanakkor a felállított modell így is előhozott érdekes tényezőket. A modell p-értéke alacsony (p-érték kisebb mint 0,01, F-érték: 2984), ami azt jelenti, van olyan magyarázóváltozónk, amely szignifikáns. A többváltozós lineáris regressziós modell becsült paramétereit az 5. táblázat mutatja.

Ha alaposabban megvizsgáljuk a magyarázóváltozókat, láthatjuk, hogy mindegyik független változó szignifikáns. A magyarázóváltozók közül az átszállások helyének (városának) nem egyezése 1% alatti szignifikanciaszinten már nem lenne szignifikáns, míg a Budapestre érkezés napszakai közül az esti érkezés 5% alatt már nem szignifikáns, ugyanakkor a modellben ezt a változót is szerepeltetem, tekintve, hogy a modellt 5%-os szignifikanciaszinten vizsgálom.

Az 5. táblázatban a regressziós együtthatók vizsgálatakor nem szabad elfelejteni, hogy a kapott érték amerikai dollárban szerepel, így a könnyebb értelmezhetőség miatt a forintban jellemző értékeket is megjelenítettem a 2023. augusztus 28-i középárfolyamon. Szürke háttérrel azokat az értékeket emeltem ki, amelyek egy, vagy akár több repülőjegy vásárlása esetén már jelentős árbeli különbséget okozhatnak. Az eredmények alapján látható, hogy több magyarázóváltozó is negatívan hat az árra.

Érdekes megfigyelés, hogy szintén jelentős különbséget okozhat az ajánlatok lekérésének a hónapon belüli időpontja. Ezt a magyarázóváltozót három részre bontottam a vizsgálat során (hónap eleje, közepe, vége), és megfigyelhető, hogy a hónap elején történő lekérdezés megdrágítja a jegyet, míg a hónap vége felé csökkenő hatás tapasztalható. A lekérdezésnél az a tényező, hogy hétközben vagy hétvégén töltöttük le az ajánlatot, nem gyakorolt nagy hatást az árra, ugyanakkor mégis szignifikáns. Gyakran lehet hallani, hogy az utazás időpontjához közeledve

egyre drágábbak lesznek a jegyárak. Ezt a tényezőt is vizsgáltam, és az adott feltevések mellett megfigyelt adatok azt mutatják, hogy ez a tényező ugyan szignifikáns, de csak csekély mértékben emeli meg az árat.

5. táblázat

A felépített regressziós modell adatai
The data of the constructed regression model

Magyarázóváltozók	Együttható	Nem robusztus standard hiba			Robusztus standard hiba			Változás forintban
		standard hiba	t-érték	p-érték	standard hiba	t-érték	p-érték	
ajánlatok szama	-0,14	0,00	-38,9	< 0,01***	0,00	-32,3	< 0,01***	-51
lesz_honaphónap_közepe	-18,61	1,01	-18,4	< 0,01***	1,00	-18,6	< 0,01***	-6 607
lesz_honaphónap vége	-21,45	1,00	-21,5	< 0,01***	1,05	-20,5	< 0,01***	-7 615
lesz_hetk_hetvhétvége_ejszaka	3,19	0,86	3,7	< 0,01***	0,85	3,8	< 0,01***	1 134
ind_nap_össz_szama	-0,85	0,02	-44,1	< 0,01***	0,02	-44,5	< 0,01***	-301
napszak_1délután	-27,19	1,09	-25,0	< 0,01***	1,15	-23,6	< 0,01***	-9 652
napszak_1hajnal	-5,59	1,16	-4,8	< 0,01***	1,24	-4,5	< 0,01***	-1 986
atsz_1_tajkiterő_nélkül	22,97	1,94	11,8	< 0,01***	2,60	8,8	< 0,01***	8 154
atsz_1_lak	0,00	0,00	7,2	< 0,01***	0,00	6,6	< 0,01***	0
ind_1_hetk_hetvhétvége	12,67	1,00	12,7	< 0,01***	12,67	12,8	< 0,01***	4 498
oda_utaz_ido_e	0,01	0,00	55,3	< 0,01***	0,00	51,3	< 0,01***	2
erk_2_napszakkdélután	-17,76	1,23	-14,4	< 0,01***	1,24	-14,4	< 0,01***	-6 305
ind_3_hetk_hetvhétvége	-12,65	0,88	-14,3	< 0,01***	0,90	-14,1	< 0,01***	-4 491
ind_3_napszakeste	19,23	1,16	16,6	< 0,01***	0,99	19,5	< 0,01***	6 827
ind_3_napszakhajnal	84,70	3,11	27,2	< 0,01***	3,31	25,6	< 0,01***	30 069
atsz_helye_oda_visszanem_egyezik	-2,37	1,04	-2,3	0,02*	0,92	-2,6	0,01*	-841
vissza_utazasi_ido_e_arfolyam	0,00	0,00	51,9	< 0,01***	0,00	46,6	< 0,01***	1
atsz_legit_1_4nem_egyezik	-1,05	0,06	-16,3	< 0,01***	0,06	-16,2	< 0,01***	-371
atsz_legit_1_4nem_egyezik	-47,37	1,79	-26,5	< 0,01***	2,10	-22,6	< 0,01***	-16 816
atsz_legit_2_3nem_egyezik	41,21	2,12	19,4	< 0,01***	2,35	17,6	< 0,01***	14 630
atsz_legit_össz_nem_egyezik	123,40	1,86	66,2	< 0,01***	1,69	73,0	< 0,01***	43 807
nyarvár_előtt	94,99	1,63	58,3	< 0,01***	1,62	58,6	< 0,01***	33 721
ind_3_tájékoztató_nélkül	-13,71	1,99	-6,9	< 0,01***	2,50	-5,5	< 0,01***	-4 867
ind_3_lakosság	0,00	0,00	34,7	< 0,01***	0,00	29,9	< 0,01***	1
bp_erk_napszakkdélután	7,64	1,14	6,7	< 0,01***	1,19	6,4	< 0,01***	2 714
bp_erk_napszakeste	6,75	2,33	2,9	< 0,01**	2,48	2,7	< 0,01**	2 397

Megjegyzés: a táblázatban a * a $p < 0,1$, a ** a $p < 0,005$, a *** a $p < 0,001$ szintű szignifikanciát jelöli.
Forrás: saját szerkesztés.

A Budapestről való indulás fontos tényező, mivel ha délelőtt indulunk, akkor napközben több átszállási lehetőség áll rendelkezésre, ami azt eredményezi, hogy gyorsabban jutunk el a célunkhoz, valamint feltételezhető az is, hogy még aznap megérkezünk, arról nem is beszélve, hogy a délelőtti indulás nekünk utazóknak is kényelmesebb, gondoljunk csak például a reptérre való kijutásra tömegközlekedéssel. A modellben ennek a tényezőnek is nagy hatása van, ugyanis a délelőtti induláshoz képest a délutánira átlagosan több mint 9600 forinttal olcsóbb repülőjegyet kapunk, hajnali indulásnál pedig ez átlagosan 1986 forintot jelent. Érdekes jobban megvizsgálni a New Yorkból induló járatokhoz tartozó magyarázóváltozókat is. A New Yorkból induló járat indulási időpontja (napszaka) jelentős árbeli hatással bír, ugyanis a délutáni járatokhoz képest a hajnali átlagosan 30 069 forinttal többet kerül. Hasonlóan nagy mértékű különbséggel szembesülünk akkor is, ha azt vizsgáljuk, hogy a légitársaságok minden útvonalon megegyeznek-e. Amikor ugyanis a négy repülőút során a légitársaság nem ugyanaz, átlagosan 43 807 forinttal drágább a repülőjegy. Fontos ennél a pontnál megjegyezni, hogy a légitársaságok szövetségeit nem választottuk külön, azaz csak a légitársaságok azonosságát néztük. A magyarázóváltozók közül érdemes még egyet kiemelni az eredmények alapján. Az ajánlatok lekérésének olyan szempontból is jelentős szerepe van, hogy az ajánlatot az arra nyáron, vagy még nyár előtt töltöttük le. Meglepő, de a vizsgálat során arra a következtetésre jutottam, hogy a nyár előtt lekérdezett ajánlatok árai átlagosan 33 721 forinttal drágábbak, mint a nyáron lekérdezettek. Ez betudható akár annak, hogy nyáron több foglalás történik, de akár annak is, hogy a nyaralások miatt sűrűbbek azok a járatok, amelyek az átszállások helyeit adják.

3.2. Interakciók beépítése a modellbe

A modell alaposabb vizsgálatához interakciókat is számoltam, illetve ábrázoltam. Erre azért volt szükség, mert a magyarázóváltozók között lehetnek interakcióban lévő változók, azaz az egyik magyarázóváltozó marginális hatásának nagyságát befolyásolhatja a másik szintje. Ilyen interakció lehet az ajánlatok és az indulás előtti napok száma, mivel ebben az esetben feltételezhető, hogy minél inkább csökken az indulás előtti idő, annál kisebb az ajánlatok hatása, azaz az indulás előtti napok száma sztochasztikusan hat az ajánlatok számának hatására is. Interakciókat építettem be az alábbi tényezők közé:

- az ajánlatok száma és az indulás előtti napok száma,
- a Budapestről New Yorkba történő indulás ideje (napszak) és a New Yorkból Budapestre történő indulás ideje (napszak).

A többváltozós lineáris regressziós modellt és az interakcióval bővített többváltozós lineáris regressziós modellt Wald-próbával teszteltem. A teszt alapján a p-érték kisebb mint 0,01.

Az interakciók beépítése után a modell determinációs együtthatója megnövekedett, 0,406 lett, ami azt jelenti, hogy a magyarázóváltozók együttese az interakciók segítségével 40,6%-ban magyarázzák az árak varianciáit.

6. táblázat

Az interakcióval kibővített regressziós modell adatai
Data of the interaction expanded regression modell

Magyarázóváltozók	Együttható	Nem robusztus standard hiba			Változás forintban
		standard hiba	t-érték	p-érték	
ajánlatok_szama	0,91	0,02	54,90	< 0,01***	324
ind_nap_ossz_szama	4,65	0,09	53,73	< 0,01***	1 652
lesz_honaphónap_közepe	-22,05	0,99	-22,25	-22,246	-7 828
lesz_honaphónap_vége	-22,85	0,98	-23,42	< 0,01***	-8 112
lesz_hetk_hetvhétvége	1,03	0,84	1,23	0,22	366
ejszaka	0,77	0,16	4,93	< 0,01***	273
napszak_1délután	-49,86	2,06	-24,26	< 0,01***	-17 700
napszak_1hajnal	-46,65	2,18	-21,42	< 0,01***	-16 561
ind_3_napszakeste	-10,07	2,15	-4,69	< 0,01***	-3 575
ind_3_napszakhajnal	-4,28	4,58	-0,94	0,35	-1 518
atsz_1_tajkiterő_nélkül	27,58	1,91	14,46	< 0,01***	9 791
atsz_1_lak	0,00	0,00	8,35	< 0,01***	0
ind_1_hetk_hetvhétvége	12,31	0,98	12,59	< 0,01***	4 370
oda_utaz_ido_e	22,86	0,38	59,38	< 0,01***	8 115
vissza_utazasi_ido_e	12,56	0,23	54,38	< 0,01***	4 459
erk_2_napszakedélután	-23,41	1,23	-19,02	< 0,01***	-8 311
ind_3_hetk_hetvhétvége	-14,01	0,86	-16,23	< 0,01***	-4 974
atsz_helye_oda_visszanem_egyezik	-6,74	1,03	-6,52	< 0,01***	-2 392
arfolyam	-1,55	0,06	-24,50	< 0,01***	-549
atsz_legit_1_4nem_egyezik	-44,76	1,77	-25,30	< 0,01***	-15 890
atsz_legit_2_3nem_egyezik	42,38	2,09	20,31	< 0,01***	15 045
atsz_legit_ossznem_egyezik	118,90	1,85	64,40	< 0,01***	42 210
nyaryár_előtt	100,60	1,60	62,91	< 0,01***	35 713
ind_3_tájéltáskítérő_nélkül	-14,09	1,95	-7,23	< 0,01***	-5 002
ind_3_lakosság	0,00	0,00	36,16	< 0,01***	1
bp_erk_napszakedélután	9,49	1,13	8,41	< 0,01***	3 368
bp_erk_napszakeste	2,96	2,29	1,29	0,20	1 050
ajánlatok_szama:ind_nap_ossz_szama	-0,01	0,00	-64,97	< 0,01***	-2
napszak_1délután:ind_3_napszakeste	28,64	2,37	12,09	< 0,01***	10 167
napszak_1hajnal:ind_3_napszakeste	48,52	2,53	19,16	< 0,01***	17 225
napszak_1délután:ind_3_napszakhajnal	76,81	5,09	15,09	< 0,01***	27 268
napszak_1hajnal:ind_3_napszakhajnal	220,90	6,06	36,44	< 0,01***	78 420

Megjegyzés: a táblázatban a * a $p < 0,1$, a ** a $p < 0,005$, a *** a $p < 0,001$ szintű szignifikanciát jelöli.

Forrás: saját szerkesztés.

A magyarázóváltozók együtthatóit áttekintve láthatjuk (6. táblázat), hogy ha eggyel több ajánlatot kapunk, akkor az indulás előtti napok számának az árra gyakorolt hatása átlagosan 2 forinttal csökken, ami csekély mértékű. Az induláshoz kapcsolódó interakciók ellenben jelentősebb mértékű hatást mutatnak. Ha hajnalban indulunk Budapestről, az a New Yorkból induló hajnali járat idejének árra gyakorolt hatását már átlagosan 78 420 forinttal megnöveli.

4. Következtetések

Tanulmányom célja az volt, hogy bemutassa, mely tényezők hogyan voltak hatással a repülőjegyek árára a vizsgált időszakban a Budapest–New York útvonalon. A mindennap lekért adatok elegendő információval szolgáltak ahhoz, hogy bizonyítsam, a vizsgált 23 magyarázóváltozó mindegyike jelentős hatással volt az árra, ugyanakkor az is egyértelműen látszik, hogy csupán a légitársaságok egyezősége, az indulások ideje (napszaka), illetve a lekérdezés (foglalás) ideje csökkentette, illetve növelte az árat számottevő mértékben.

A szakirodalmi áttekintés alapján is egyértelmű, hogy számos olyan tényező van egy-egy ár meghatározásakor, amit a *webscraping* módszerével nem feltétlenül tudunk letölteni, például a repülőgépek telítettsége, az árdiszkriminációk, illetve mivel dinamikus árazásról beszélünk a repülőjegyek esetében, így ennek nyomon követése is bonyolult feladat. Ennek ellenére a modellem bebizonyította, hogy még a fogyasztók által elérhető információk is segíthetnek abban, hogy képet kapjunk az árazásról. Az előrejelzések komplexek, ezért teljes mértékben nem is tudunk jó előrejelzést adni, de a *machine learning* módszere sokat segíthet a modellek pontosabbá tételében, ami további elemzési lehetőséget jelent a téma vizsgálatára és bővítésére.

Irodalom

- Abayi, M. – Khoshtinat, B. (2016): Study of the Impact of Advertising on Online Shopping Tendency for Airline Tickets by Considering Motivational Factors and Emotional Factors. *Procedia Economics and Finance*, 36, 532–539. [https://doi.org/10.1016/S2212-5671\(16\)30065-X](https://doi.org/10.1016/S2212-5671(16)30065-X)
- Abdella, J. A. – Zaki, N. M. – Shuaib, K. – Khan, F. (2021): Airline ticket price and demand prediction: A survey. *Journal of King Saud University – Computer and Information Sciences*, 33(4), 375–391. <https://doi.org/10.1016/j.jksuci.2019.02.001>

- Alderighi, M. – Gaggero, A. A. – Piga, C. A. (2022): Hidden prices with fixed inventory: Evidence from the airline industry. *Transportation Research Part B: Methodological*, 157(C), 42–61. <https://doi.org/10.1016/j.trb.2022.01.001>
- Aslani, S. – Modarres, M. – Sibdari, S. (2014): On the fairness of airlines' ticket pricing as a result of revenue management techniques. *Journal of Air Transport Management*, 40(C), 56–64. <https://doi.org/10.1016/j.jairtraman.2014.05.004>
- Dudás G. – Boros L. – Pál V. (2016): A repülőjegyárak változásának időbeli alakulása, budapesti indulással. *Területi Statisztika*, 56(6), 660–680. <https://doi.org/10.15196/TS560604>
- Dudás G. – Boros L. – Pál V. – Pernyész P. (2016): A költségtávolság alakulása a diszkont és a hagyományos légitársaságok repülőjegyjárai alapján, Budapest példáján. *Tér és Társadalom*, 30(2), 107–132. <https://doi.org/10.17649/TET.30.2.2732>
- Fiore, A. – Zanzalari, D. (2023): The competitive effects of product unbundling: Evidence from U.S. airlines. *Journal of Air Transport Management*, 112, 102454. <https://doi.org/10.1016/j.jairtraman.2023.102454>
- Gaggero A. A. – Luttmann, A. (2023): The determinants of hidden-city ticketing: Competition, hub-and-spoke networks, and advance-purchase requirements. *Transportation Research Part E: Logistics and Transportation Review*, 173, 103086. <https://doi.org/10.1016/j.tre.2023.103086>
- Financial Times (2023): *Air fares soar above inflation as carriers cash in on travel demand.* <https://www.ft.com/content/223e693b-caa6-48e5-8dd4-44302d068f96>
- Hunyadi L. (2006): A heteroszkedaszticitásról egyszerűbben. *Statisztikai Szemle*, 84(1), 75–82.
- Ivaldi, M. – Petrova, M. – Urdanoz, M. (2021): Airline cooperation effects on airfare distribution: An auction-model-based approach. *Transport Policy*, 115, 239–250. <https://doi.org/10.1016/j.tranpol.2021.11.006>
- Kerkemehos, Y. – Pennings, E. – Karreman B. – van Reeve, P. (2023): Price asymmetries and the path dependence of market power: Evidence from the U.S. airline industry. *International Journal of Industrial Organization*, 87, 102921. <https://doi.org/10.1016/j.ijindorg.2023.102921>
- Koc, I. – Arslan, E. (2021): Dynamic ticket pricing of airlines using variant batch size interpretable multi-variable long short-term memory. *Expert Systems with Applications*, 175, 114794. <https://doi.org/10.1016/j.eswa.2021.114794>
- Koo, B. – Mantin, B. – O'Connor, P. (2011): Online distribution of airline tickets: Should airlines adopt a single or a multi-channel approach? *Tourism Management*, 32(1), 69–74. <https://doi.org/10.1016/j.tourman.2009.11.008>
- Kovács P. (2008): A multikollinearitás vizsgálata lineáris regressziós modellekben. *Statisztikai Szemle*, 86(1), 38–67.
- Kovács P. – Petres T. – Tóth L. (2004): Adatállományok redundanciájának mérése. *Statisztikai Szemle*, 82(6–7), 595–604.
- KSH Stadat (2023): *24.1.1.19. Összefoglaló adatok a Budapest Liszt Ferenc Nemzetközi Repülőtér forgalmáról.* https://www.ksh.hu/stadat_files/sza/hu/sza0019.html (Letöltés ideje: 2023. augusztus 28.)
- Law, R. – Leung, R. – Deniczi Guillet, B. – Lee, H. A. (2011): Temporal changes of airfares toward fixed departure date. *Journal of Travel and Tourism Marketing*, 28(6), 615–628. <https://doi.org/10.1080/10548408.2011.598740>
- Lewis, M. S. (2021): Identifying airline price discrimination and the effect of competition. *International Journal of Industrial Organization*, 78, 102761. <https://doi.org/10.1016/j.ijindorg.2021.102761>

- Martínez, M. A. – Navarro, J. A. – Trinquescoste, J. (2017): The effect of destination type and travel period on the behavior of the price of airline tickets. *Research in Transportation Economics*, 62, 37–43. <https://doi.org/10.1016/j.retrec.2017.03.003>
- Morlotti, C. – Redondi, R. (2023): The impact of COVID-19 on airlines' price curves. *Journal of Air Transport Management*, 107, 102328. <https://doi.org/10.1016/j.jairtraman.2022.102328>
- Oh, J. – Huh, W. T. (2022): Hidden city travel and its impact on airfare: The case with competing airlines. *Transportation Research Part B: Methodological*, 156, 101–109. <https://doi.org/10.1016/j.trb.2021.12.006>
- Ruiz-Mafe, C. – Sanz-Blas, S. – Hernandez-Ortega, B. – Brethouwer, M. (2013): Key drivers of consumer purchase of airline tickets: A cross-cultural analysis. *Journal of Air Transport Management*, 27, 11–14. <https://doi.org/10.1016/j.jairtraman.2012.10.010>
- Sampaio, R. M. B. – Urdanoz, M. (2022): Airlines' cooperation in the domestic market: Measuring the evolution of price gaps. *Transportation Research Part A: Policy and Practice*, 166, 424–443. <https://doi.org/10.1016/j.tra.2022.11.007>
- Silva, T. C. – Dias, F. A. M. – Dos Reis, V. E. (2022): The role of network topology in competition and ticket pricing in air transportation: Evidence from Brazil. *Physica A: Statistical Mechanics and its Applications*, 601(1), 127602. <https://doi.org/10.1016/j.physa.2022.127602>
- Yu, S. (2008): Price perception of online airline ticket shoppers. *Journal of Air Transport Management*, 14(2), 66–69. <https://doi.org/10.1016/j.jairtraman.2007.11.002>
- Világgazdaság (2020): Új árazási stratégia a légitársaságoknál. <https://www.vg.hu/cegvilag/2020/11/uj-arazasi-strategia-a-legitarsasagoknal>
- Wang, Z. – Ye, Y. (2012): Hidden-city ticketing: The cause and impact. *Transportation Science*, 50(1), 1–26. <https://doi.org/10.1287/trsc.2015.0587>