

OTDK - DOLGOZAT

2025



HOGYAN NYERD MEG A PREMIER LEAGUE-ET?

HOW TO WIN THE PREMIER LEAGUE?

AVAGY MELY SPORTSZAKMAI ÉS GAZDASÁGI MUTATÓSZÁMOK JELZIK
ELŐRE A CSAPATOK HOSSZÚ TÁVÚ SIKEREIT?

Készítette: Ferencz Márton

Budapesti Corvinus Egyetem,
Gazdálkodási és menedzsment BSc,

Témavezető: Madari Zoltán

Kézirat lezárása: 2024. december 3.

Absztrakt

A dolgozat célja, hogy megvizsgálja a futballcsapatok teljesítményére ható legfontosabb gazdasági és sportteljesítmény-mutatókat. Ehhez elsőként irodalmi áttekintést végez a releváns sportgazdaságtani és empirikus szakirodalom bemutatásával. A dolgozat panel ökonometriai elemzési módszertant alkalmazva vizsgálja az elmúlt hét teljes szezon adatait az angol labdarúgó első osztályban, melyek alapjául az Fbref, Capology és Transfermarkt adatbázisai szolgálnak. Először egy alap Pooled OLS modellt épít, melyben az eredményváltozó az elért pontszám és ezt próbálja magyarázni 26 másik változó felhasználásával. Majd ezt a modellt specifikálja két új modellre; az egyik a hagyományos mutatókat használja, a másik a haladó statisztikákra alapoz. Mindegyik modellen elvégéz modellszelekciót, illetve standardizálja is azok folytonos változóit az összehasonlíthatóság érdekében. A kapott eredmények alapján a dolgozat megállapítja, hogy a szakirodalomhoz hasonlóan a lőtt és kapott gólok hatása jelentős, a gazdasági mutatók közül a csapatok teljes piaci játékosértéke a legszignifikánsabb, és a haladó statisztikák közül a várható gólkülönbségnek van a teljesítményre a legnagyobb hatása.

Journal of Economic Literature (JEL) kód: L83, C23, M21

Kulcsszavak: Labdarúgás; Panel adatmodell; Teljesítménymutatók; Gazdasági hatás; Haladó statisztikai mutatók

Tartalom

1.	Bevezetés.....	1
2.	Szakirodalmi összefoglaló.....	2
2.1	Üzleti ligák és működésük	2
2.1.1	Bevételek.....	3
2.1.2	Kiadások.....	6
2.1.3	Transzferpiac	7
2.2	A Premier League létrejötte.....	10
2.3	Empirikus szakirodalom.....	11
2.3.1	A kapcsolat és aktualitása.....	11
2.3.2	A modern adatelemzés a labdarúgásban.....	12
2.4	Teljesítményre ható mutatószámok a szakirodalomból.....	13
2.4.1	Gazdasági hatás	13
2.4.2	Sportszakmai hatás	15
3.	Módszertan	19
4.	Adatbázis bemutatása	22
5.	Elemzés	25
5.1	Leíró statisztika	25
5.2	Panel elemzés	26
5.2.1	Alapmodell	26
5.2.2	Specifikus modellek	30
6.	Eredmények.....	33
7.	Kutatási korlátok	35
8.	Kitekintés, továbbfejlesztési lehetőségek.....	36
9.	Összefoglalás.....	36
10.	Forrásjegyzék	39
11.	Mellékletek.....	45

Ábrajegyzék

1. ábra: Premier League klubok bevételei 2020/21-2024/25 - Forrás: Saját szerkesztés Deloitte (2024) alapján.....	5
2. ábra: Premier League játékosok átlagos fizetése klubonként 2022/23 millió GBP-ben - Forrás: Saját szerkesztés Statista (2023) alapján.....	6
3. ábra: Folytonos változók hisztogram ábrái - Forrás: Saját szerkesztés.....	26
4. ábra: Heteroszkedaszticitás ábrázolása - Forrás: Saját szerkesztés.....	27

Táblázatjegyzék

1. táblázat: Változók bemutatása - Forrás: Saját szerkesztés	23
2. táblázat: Szűkített modell eredményei robosztus hibataggal - Forrás: Saját szerkesztés.....	29
3. táblázat: Teszt eredmények modellválasztáshoz - Forrás: Saját szerkesztés.....	31
4. táblázat: A szűkített hagyományos modell eredményei robosztus hibataggal - Forrás: Saját szerkesztés	32
5. táblázat: Leíró statisztikai elemzés eredménye a változókra - Forrás: Saját szerkesztés.....	45
6. táblázat: Az alapmodell eredményei robosztus hibataggal - Forrás: Saját szerkesztés.....	46
7. táblázat: Standardizált alapmodell eredményei - Forrás: Saját szerkesztés	47
8. táblázat: Hagományos modell eredményei robosztus hibataggal - Forrás: Saját szerkesztés	47
9. táblázat: Haladó modell eredményei robosztus hibataggal - Forrás: Saját szerkesztés	48
10. táblázat: Standardizált hagyományos modell eredményei - Forrás: Saját szerkesztés.....	48
11. táblázat: Standardizált haladó modell eredményei - Forrás: Saját szerkesztés	49

1. Bevezetés

A sport iránt érdeklődő közönség gyakran belefuthat a social médiában olyan véleményekbe, hogy egy-egy edzőnek, játékosnak semmi köze az adott sporthoz és annál bárki jobb teljesítményt tudna hozni a mérkőzéseken, mint azok, akik éppen a legjobbjukat próbálják nyújtani a pályán. Az ilyesfajta „fotelszurkolók” előszeretettel alkalmazzák, az eredményből elemzés módszerét, vagyis a mérkőzőképét teljesen figyelmen kívül hagyva rosszabb esetben pusztán az eredményből, kevésbé rossz esetben néhány nagyon alapnak mondható statisztikai mutatószám alapján vonnak le következtetéseket egy-egy csapatról. Kezdő sportújságíróként könnyű ilyesmi hibába esni; az embernek nincsen ideje megnézni minden meccset és így elemzést írnia; ilyenkor csak ránéz az ember a labdabirtoklásra mondjuk, ami azt mutatja, hogy az egyik csapat 65%-ban birtokolta a labdát és egyből megírja, hogy micsoda dominancia volt látható a pályán. Igazából viszont pusztán ebből az információmorzsából még semmi sem következik; a pálya mely részén birtokolták a labdát, hány lövést hoztak össze, milyen minőségűek voltak ezek a lövések, mikor születtek a gólok a mérkőzésen stb. Az adatokat tökéletesen lehet használni objektív mérésekre a sportban, de ezeket nagyon fontos kontextusukban kezelni és több mutatót is vizsgálni a minél teljesebb kép eléréséhez. Azonban az adatokat leginkább alátámasztásként, bizonyításként érdemes a leginkább használni egy-egy megfogalmazott állításunkhoz, elemzésünkhöz. „Az adatok a futball megértésére vannak, és nem fordítva” (Szabó et. al, 2022, 248.o.)

Az adatalapú döntéshozatal nem mindig volt jelen a labdarúgásban, viszont a modern labdarúgás egyik nagy vívmányának ezt nevezhetjük. Aktuális példák lehetnek erre a Brighton és a Brentford csapatai, akik néhány év leforgása alatt az alsóbb osztályokból stabil Premier League résztvevőkké váltak (Naylor és Harris, 2023). Dolgozatomban a labdarúgás új adatelemzési és adatgyűjtési eszközeiről, valamint a széles körben használt statisztikai mutatók bemutatásáról és alkalmazásáról írok. Kutatásom célja annak meghatározása, hogy a modern labdarúgásban vizsgált gazdasági és sportszakmai statisztikai mutatók közül melyek állnak szoros kapcsolatban egy csapat teljesítményével. Erre egy ökonometria elemzési módszert, a panel modellezést fogom használni.

Mindazonáltal egy csapat teljesítményére nem csak a sportszakmai mutatók lehetnek hatással, fontos vizsgálnunk a klubok környezetét is, amelyet főként gazdasági szempontok alapján tudunk megtenni, hiszen erről érhetőek el nyilvános adatok. A kutatásom hipotézise szerint minél több pénzügyi erőforrást fektetnek az adott klubba a tulajdonosok, annál nagyobb

a valószínűsége, hogy a csapat sikereket ér el. Ezek az erőforrások szerteágazó területekre fordíthatóak, így a megfelelő vizsgálathoz ezeket érdemes csoportosítani és szétválasztani őket.

Az elemzésem tárgya az angol labdarúgó elsőosztály; a Premier League lesz. Az elmúlt 7 szezon adatait fogom vizsgálni, azaz a 2017/2018-as szezontól kezdve a 2023/2024-es szezonig. A sportszakmai mutatókat az Fbref (2024) oldaláról használom fel, amely az OPTA adatait használja, míg a gazdasági mutatóknál több forrásból is hozok információkat; az átigazolásokkal és játékos értékekkel kapcsolatos értékeket a Transfermarkt-ról (2024), a pénzügyi mutatókat pedig a Capology (2024) oldaláról használom fel. Az adataimnak egyszerre van keresztmetszeti és idősoros jellege is, így panel modellezést fogok használni az elemzésnél. Az ökonometria elemzést R-ben végzem.

Előzetes feltevéseim között van a már fentebb említetten túl, hogy valamilyen mértékben az átigazolásokra költött összegek is befolyásolják egy csapat sikerességét, hiszen természetesnek tűnik, hogy egy drágább játékos jobbnak is nevezhető és ezzel jobban tudja elősegíteni a csapata sikerességét. Kutatásom újdonsága a haladó statisztikai mutatók használata, illetve ezek ötvözése pénzügyi, gazdasági mutatókkal. Előbbiek használata a szakirodalomban még nem vált általánossá; célom, hogy bizonyítsam ezek használhatóságát és relevanciáját.

Az eredményeim pedig azt mutatják, hogy a szakirodalomban is használt mutatók egy részének hatását sikerült nekem is reprodukálnom; ilyen változók voltak például a lőtt és kapott gólok, de sikerült újdonságot is behoznom; gazdasági téren a teljes piaci játékosérték bizonyult a legszignifikánsabb magyarázó változónak, a haladó statisztikák közül, pedig a várható gólkülönbség számít a legmeghatározóbb változónak.

Az adatelemzés a sportban túlmutat a klubok és szakemberek munkáján: az átlagos szurkolók és sportújságírók számára is értékes eszközzé vált. Részletes statisztikák segítségével a rajongók jobban megérthetik kedvenc csapatuk teljesítményét, a különböző mutatók alapján pedig megjósolhatják az elkövetkező meccsek kimenetelét. Az elemzések révén mélyebb betekintést nyerhetünk a sportágak működésébe, összehasonlíthatjuk a játékosokat és csapatokat, és objektívebben elemezhetjük a sportteljesítményt, mint valaha.

2. Szakirodalmi összefoglaló

2.1 Üzleti ligák és működésük

A sportban elért sikert két szempontból vizsgálhatjuk: sportszakmai és gazdasági-pénzügyi vonatkozásban (András, 2004). A sportszakmai siker egy klub esetében akkor következik be, ha eléri vagy meghaladja a kitűzött célokat, legyen az egy kupában vagy

bajnokságban elért helyezések, vagy más sportszakmai teljesítmény, mint például utánpótlásfejlesztés vagy játékosok fejlesztése. A pénzügyi siker azonban kevésbé jellemző a klubokra, hiszen a tulajdonosok általában nem azért fektetnek be egy csapatba, hogy nyereséget termeljenek vele. Inkább arról van szó, hogy reklámot generáljanak más vállalkozások számára, vagy esetleg klubszeretetet vezérli őket. Ha egy klub egy adott idényben több nyereséget termel, mint átlagosan, általában visszaforgatja azt a csapatba, például új játékosok vásárlására vagy infrastruktúrafejlesztésekre (András, 2004).

Az üzleti alapú sportligákban a piaci koordinációs mechanizmus dominál, ami azt jelenti, hogy a ligák igyekeznek piaci alapú bevételt szerezni, és minél kevésbé támaszkodnak állami támogatásokra vagy tagdíj jellegű befizetésekre (András és Kozma, 2015). A klubokat segíti ebben a viszonylag stabil médiaérdeklődés, a szurkolók hűsége és a szabályozás bevezetése, ami lehetővé teszi számukra a hivatásos és üzleti jellegű működést, valamint a költségek hatékony kezelését (Bács, 2020a). A sportági szakszövetségek, mint például az Európai Labdarúgó-szövetség (továbbiakban: UEFA), Nemzetközi Labdarúgó-szövetség (továbbiakban: FIFA), Magyar Labdarúgó Szövetség (továbbiakban: MLSZ) vagy az Angol Labdarúgó-szövetség (továbbiakban: FA) fontos szereplői ennek a koordinációnak, és feladataikat és szerepüket később részletesen fogom tárgyalni.

Az ilyen típusú ligák manapság a legelterjedtebbek, de talán inkább a legnézettebbeknek nevezhetők. Ebbe a kategóriába tartoznak azok a ligák, amelyek a legjobb játékosokat és edzőket vonzzák, mivel itt keresik meg a legtöbb pénzt a munkájukért. Ennek következtében magasabb színvonalú és látványosabb játék látható, ami vonzza a nézőket, és ezzel jelentős bevételt generál a ligák és a klubok számára (András és Kozma, 2015).

2.1.1 Bevételek

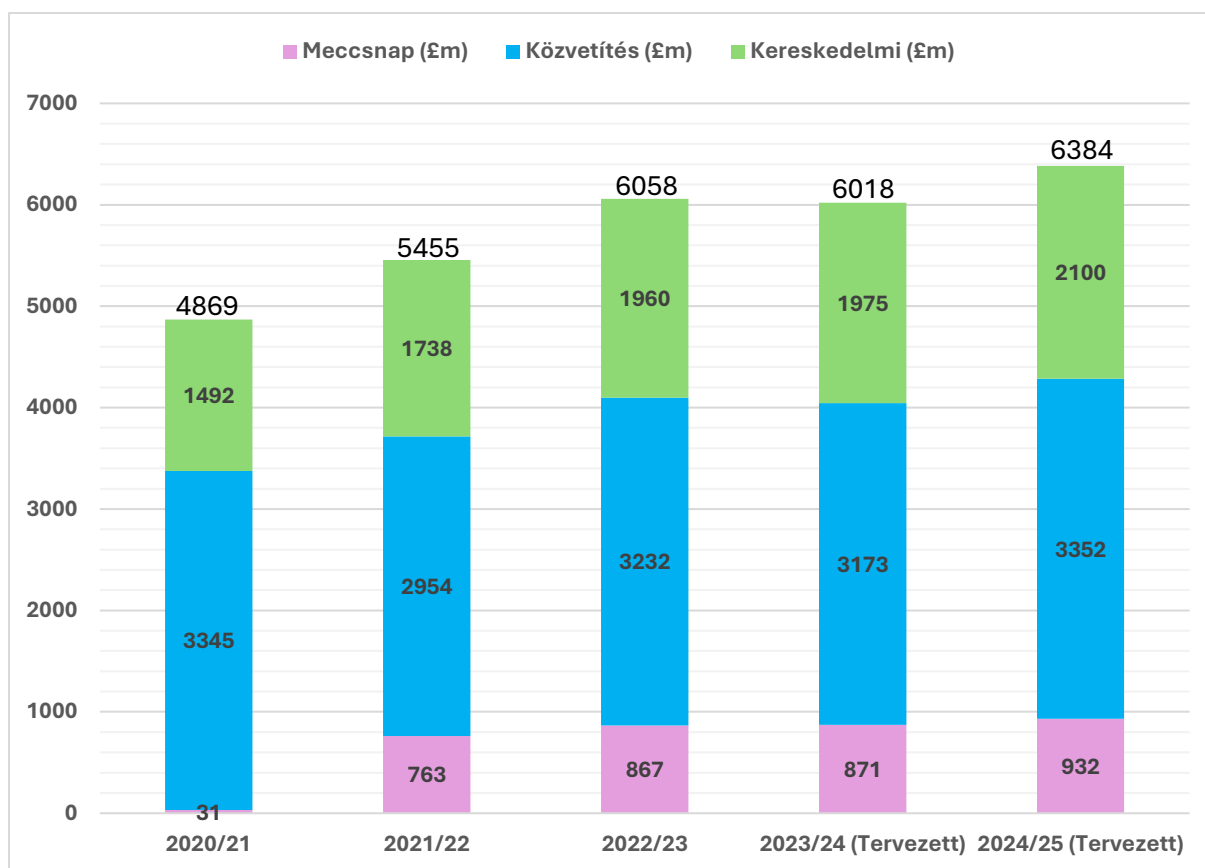
A bevételeket az UEFA 2024-es pénzügyi jelentése alapján a tevékenységhez kapcsolódóan tagolja. Az alábbiakban kategorizálták a bevételeket: a jegybevételek a meccsjegyek és bérletek eladásából származnak; a szponzori szerződések és hirdetések a főszponzoroktól és reklámokból befolyó bevételeket jelentik; a televíziós közvetítési jogok esetében a liga közvetítési bevételei és annak elosztása bajnokságtól függ; a támogatások, például az UEFA-nál érkező bevételeket jelentik; a működési és egyéb bevételek közé tartoznak a rendkívüli bevételek, valamint a merchandising termékek eladása; végül az átigazolási díjak, amelyeket az UEFA nem számít bele a klub bevételeibe, de az összefüggés kedvéért itt is szerepelnek, ezek a bruttó transzferbevételek, amelyeket a klubok elszámolásaikban külön jelentenek.

A futballklubok számára az egyik legfontosabb bevételi forrást a jegybevételek jelentik, amelyek magukban foglalják a meccsjegyek és a bérletek eladásából származó bevételeket. A szurkolók jelenléte a stadionban nemcsak a csapat számára nyújt támogatást, hanem jelentős pénzügyi hozzájárulást is biztosít. A jegybevételek mértéke nagyban függ a csapat népszerűségétől, a stadion befogadóképességétől, valamint a csapat teljesítményétől. Ezek a bevételek nem csak a hazai bajnoki mérkőzésekből származnak, hanem a nemzetközi kupákban való részvétel során eladott jegyek is jelentős összeget hozhatnak (UEFA, 2024).

A szponzorokat vizsgálva; a főszponzorok - akik gyakran a csapat mezén is megjelennek - a legfontosabbak egy klub életében, ők komoly összegeket fizetnek azért, hogy márkájuk a futballklubhoz kapcsolódjon. Ezen kívül a stadionokban elhelyezett reklámok és hirdetések is folyamatos bevételi forrást jelentenek. A különböző bajnokságok televíziós közvetítési jogai jelentős összegeket hoznak, amelyeket a liga eloszt a csapatok között. A nagyobb bajnokságok, mint például az angol Premier League vagy a spanyol La Liga, különösen magas összegeket generálnak ezen a területen, de a bevételek allokációja teljesen eltér. Ezek a bevételek a klubok pénzügyi stabilitását is erősítik, hiszen hosszú távú szerződésekről van szó, amelyek előre kiszámítható bevételt biztosítanak.

A különböző támogatások, például az UEFA-tól kapott pénzek, szintén fontos szerepet játszanak a klubok pénzügyi életében. Az UEFA különböző versenyek, mint például a Bajnokok Ligája, Európa Liga és Európa Konferencia Liga révén oszt ki jelentős összegeket a résztvevő csapatok között. Ezek a támogatások nemcsak a csapatok versenyképességét növelik, hanem a klubok költségvetését is jelentősen javítják. A működési és egyéb bevételek közé tartoznak a rendkívüli bevételek és a merchandising termékek eladásából származó összegek. A klubok saját márkájukat erősítve különféle termékeket, mint például mezeket, sálakat, és egyéb ajándéktárgyakat értékesítenek, ami jelentős bevételt generálhat. Emellett különböző rendezvények, stadiontúrák, és egyéb szolgáltatások is hozzájárulnak a klub bevételeihez. Bár az UEFA az átigazolási díjakat nem számítja bele a klub bevételeibe, mégis fontos részét képezik a klubok pénzügyi stratégiájának. Az átigazolási díjakból származó bevételek, amelyeket a játékosok eladásából szereznek, jelentős összegeket jelenthetnek. Ezek az összegek a klubok költségvetésében külön szerepelnek, és gyakran visszaforgatják őket új játékosok szerződtetésére, vagy egyéb klubfejlesztési célokra (UEFA, 2024). A transferpiacról még egy későbbi fejezetben lesz szó.

A Deloitte (2024) elemzéseiben a bevételek minimálisan eltérő struktúrában vannak. Náluk 3 kategóriára oszlanak a bevételek; meccsnapi bevételek, közvetítési bevételek, illetve kereskedelmi bevételek.



1. ábra: Premier League klubok bevételei 2020/21-2024/25 - Forrás: Saját szerkesztés Deloitte (2024) alapján

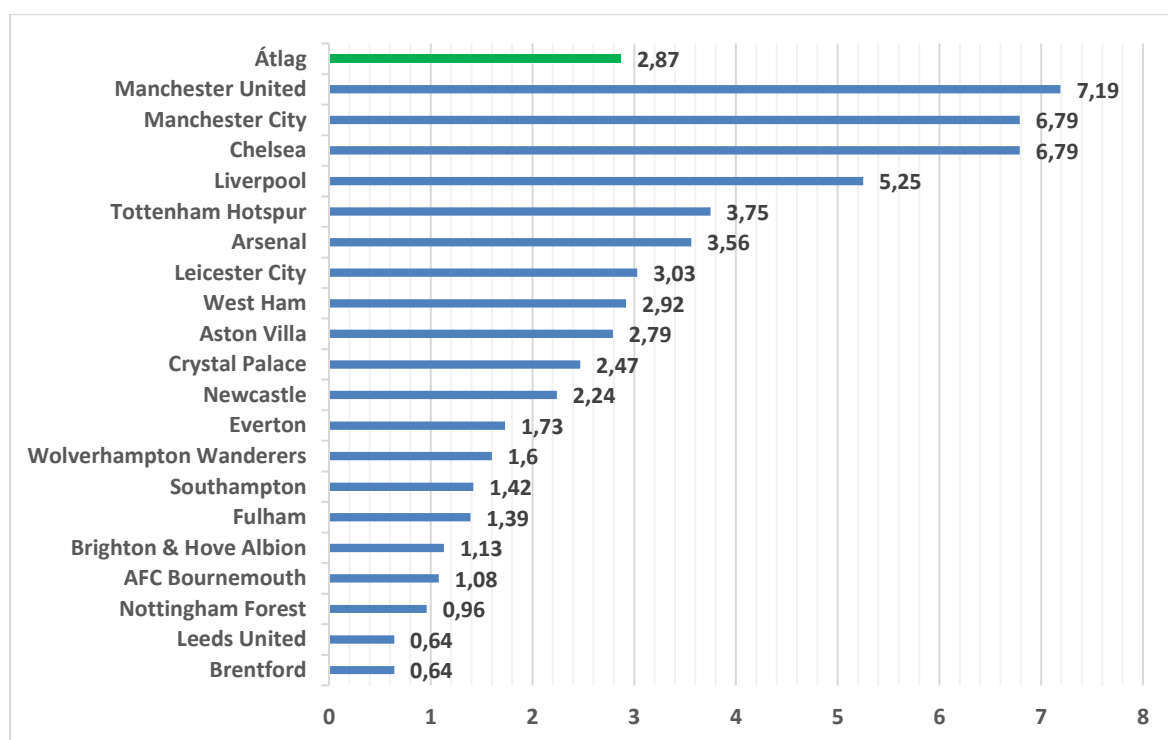
A Deloitte jelentése alapján a Premier League klubjai a 2022/23-as szezonban rekordbevételt értek el, összesen 6,058 milliárd fontot, ami 11%-os növekedést jelent az előző évhez képest. A „Big six” csapatok (hagyományosan a liga 6 legnagyobb és legjobban kiemelkedő klubját nevezik így; Manchester United, Manchester City, Liverpool, Arsenal, Chelsea, Tottenham), mind meghaladták a átlagos bevételt, amely 303 millió font volt, míg a többi 14 klub átlaga 184 millió font volt. Ahogyan az az 1. ábráról is leolvasható a meccsnapok bevételei 14%-kal, 867 millió fontra nőttek, köszönhetően a szinte teljesen megtelt stadionoknak és a jegyárak emelésének. A kereskedelmi bevételek is új csúcst értek el, 221 millió fonttal növekedve, közel 2 milliárd fontra. A közvetítési bevételek 9%-kal, 3,2 milliárd fontra emelkedtek, főként a Premier League új nemzetközi közvetítési jogainak köszönhetően. Az UEFA klubversenyeken részt vevő angol klubok összesen 391 millió fontot kaptak az UEFA-tól.

Általánosságban megjegyezhető, hogy a kategóriákon belül jelentős különbségek mutatkoznak a bajnokságok és akár a csapatok között is. A nagyobb, globálisabb klubok saját márkájú termékei iránt nagyobb a kereslet, mint a kisebb, közép-méretű csapatokéra. Ez a nézettségre is érvényes, azonban a szövetségek igyekeznek kiegyensúlyozni a televíziós

közvetítési jogokból származó bevételt, hogy minden csapatnak jusson belőle egy megfelelő rész. Mindazonáltal a lényeg az, hogy minden klub esetében az üzleti sikeresség közvetlenül összefügg a sportteljesítménnyel (Bács, 2020b).

2.1.2 Kiadások

A teljesség kedvéért nézzük meg az „egyenlet” másik oldalát is; a bevételek után a ligák kiadásait, ugyancsak a már fentebb is említett UEFA jelentés alapján. Ezeket a következőképpen lehet csoportosítani: alkalmazottak fizetése (játékosok és személyzet is), játékosokért kifizetett átigazolási díjak és a működési kiadások (meccsnap kiadásai, reklám kiadások, egyéb működési kiadások), beruházások.



2. ábra: Premier League játékosok átlagos fizetése klubonként 2022/23 millió GBP-ben - Forrás: Saját szerkesztés Statista (2023) alapján

A már korábban is említett UEFA és Deloitte jelentéseket vizsgálva a 2022/23-as szezonban a Premier League klubok költségei és adóssága jelentős változásokon mentek keresztül, melyek különböző területeken mutattak markáns növekedést és csökkenést. A klubok összesített bérköltsége először lépte át a 4 milliárd fontos határt, amely 10%-os növekedést jelentett az előző szezonhoz képest. Átlagban egy Premier League klub egy játékosának 2,87 millió fontot fizetett ki a 2022/23-as szezonban (Statista, 2023). A klubok nettó adóssága 473 millió fonttal nőtt, így elérte a 3,1 milliárd fontot, ami főként az infrastruktúra projektek finanszírozásának köszönhető. Az Everton adóssága például 189 millió fonttal emelkedett az új

Bramley-Moore Dock Stadion építése miatt, így összesen 331 millió fontot tett ki (UEFA, 2024).

A Premier League klubok a 2024/25-ös szezonban próbajellelleggel bevezetik a költségsapkákat, amelyek célja, hogy a klubok bérköltségei, ügynöki és átigazolási költségei ne haladják meg a klub által generált bevételek 85%-át, vagy 70%-át azoknál a kluboknál, amelyek részt vesznek a UEFA klubversenyeiben. Ezeket a költségeket egy abszolút sapka is korlátozza, amely a liga legkisebb központi sugárzási és kereskedelmi bevételeinek többszöröse. Annak ellenére, hogy a Premier League klubok bevételei 226 millió fonttal növekedtek, meghaladva a bérköltségek 377 millió fontos emelkedését, a működési nyereség 18%-kal, 393 millió fontra csökkent. Ez a csökkenés főként az egyéb üzemeltetési költségek 1,6 milliárd fontra történő emelkedésének tudható be. Továbbá, a Premier League klubok immár ötödik éve jelentettek összesített adózás előtti veszteséget, amely 2022/23-ban 685 millió fontot tett ki, ami 14%-os növekedést jelent az előző évhez képest. A 2022/23-as szezonban a Premier League klubok nettó adóssága 473 millió fonttal nőtt, elérve a 3,1 milliárd fontot. (UEFA, 2024 és Deloitte, 2024)

Az összes kategóriáról megállapíthatjuk, hogy minél nagyobb egy klub, annál többet kell ezen a kategórián belül költenie (még, ha ezen megnövekedett kiadások kategóriánként nem is méretarányos). Minél nagyobb egy klub, annál nagyobb az infrastruktúrája és az elvárható szakmai fejlettség, az ehhez szükséges személyzet is nő, ha bővítjük az infrastruktúrát, illetve a szakmailag képzetesebb szakemberek is magasabb fizetést követelhetnek, többre kerül egy nagyobb stadiont fenntartani stb.

A Deloitte adatai alapján az összesített működési profit 18%-kal, 393 millió fontra csökkent, míg az adózás előtti veszteség 685 millió font volt, ami 14%-os növekedést jelent. A klubok nettó adóssága 473 millió fonttal nőtt, elérve a 3,1 milliárd fontot, főként infrastruktúra projektek finanszírozása miatt. A Premier League klubjai új pénzügyi szabályok bevezetését tervezik a költségek ellenőrzése érdekében, amelyek a 2024/25-ös szezontól kezdve próba jellelleggel lesznek érvényben. Ezek a szabályok a klubok bérköltségeit és átigazolási kiadásait a bevételek meghatározott százalékára korlátozzák, ami várhatóan segíti a fenntartható pénzügyi működést a jövőben.

2.1.3 Transzferpiac

Egy közös metszete lehet az előző két fejezetnek, vagyis ahol egyszerre szerezhetsz bevételt és keletkezhetnek kiadásai egy klubnak az a transzferpiac. A játékosok átigazolási piacának megértéséhez alapfogalmak tisztázása szükséges. „A játékospiac egy speciális

munkaerőpiac, ahol a szerződés tárgya alapvetően egy speciális vagyoni értékű személyi jog: a játékjog használatának meghatározott időtartamra vonatkozó átruházása. Olyan személyhez kötődő vagyoni értékű jog, mely a sportoló sporttevékenységhez fűződő fizikai és szellemi képességeinek összességét fejezi ki.” (András, 2004, 21.o.) Ez a jog lehet átruházható egy másik klubba, akár véglegesen, amiért átigazolási díjat kell fizetni, vagy ideiglenesen, ilyenkor a kölcsönszerződés érvényes, ami tartalmazza a játékos alapfizetését és a kölcsönvételi díjat a játékos eredeti klubjának. Az átigazolási díj egy élő szerződésű játékosért csak egy bizonyos összegért cserélhető. Ez az összeg lehet előre meghatározott a játékos szerződésében, vagy a két klubnak közösen kell megegyeznie róla. A díjat befolyásolja több tényező, mint például a játékos életkora, sérülékenysége, posztja és aktuális szerződése. A játékosok becsült piaci értékét több platform, például a Transfermarkt is megpróbálja megbecsülni. Általában a játékosokat egy játékosügynök vagy menedzser képviseli, akik közvetítőként szerepelnek a klubok és a játékosok között, miután megállapodtak az átigazolás részleteiről és a vételárról. A szerződések gyakran két részből állnak: a garantált, rendszeres jövedelemből, valamint a bónuszokból, melyek gyakran teljesítmény alapúak. Ha egy játékos szerződése lejár egy klubnál, ingyen igazolhatóvá válik, és szabadon tárgyalhat más klubokkal, átigazolási díj fizetése nélkül, amit a Bosman-szabály (Burton, 2022) biztosít 1995 óta, amikor Jean-Marc Bosman az Európai Bírósághoz fordult, miután klubjának, az RC Liège-nek lejárt a szerződése, és átigazolási díjat kértek érte. A bíróság döntése szerint ez az eljárás ellentétes a személyek szabad mozgásának jogával, ami az Európai Unió alapelve, ezért a játékosoknak nem kell átigazolási díjat fizetniük, ha szerződésük lejárt.

Az elmúlt időszakban mind az átigazolási díjak, mind a játékosok fizetései meredeken emelkedtek. Talán a legelső és legfeltűnőbb példa erre Neymar 2017-es átigazolása, amikor az új klubja, a PSG 222 millió eurót – azaz átigazolási világrekordot - fizetett a Barcelona csapatának a játékjogáért (NSO, 2017). A játékosok fizetéseinek növekedését alátámasztják az alábbi adatok is: 2023-ban az átlagos Premier League játékos éves keresete közel 3 millió dollár volt (Statista, 2023), míg 2000-ben ez az összeg 600-700 ezer dollár körül volt (Miller és Harris, 2011). Globálisan többféle módszer terjedt el ennek szabályozására; a leggyakoribb megoldás a fizetési plafon (salary cap) (András és Kozma, 2015). Az UEFA az elszálló átigazolási díjakat próbálja szabályozni a Financial Fair Play rendszerével. Ennek értelmében egy klub nem költhet többet, mint amennyi bevétele volt az előző szezonban, ezzel biztosítva a klubok fennmaradását és megelőzve a gyors csődbe jutást.

A puha költségvetési korlát jelenségét - mely fogalmat még Kornai János alkott meg – Szymanski (2022) próbálta megmagyarázni. Ezen elmélet szerint a futballklubok pénzügyi

fegyelme lazább, mert a klubokat gyakran "mentőövekkel" támogatják meg, például tulajdonosi befektetésekkel vagy közösségi forrásokkal, amikor fizetéseképtelenné válnak. Emiatt a klubvezetés kevésbé szembesül közvetlen pénzügyi következményekkel, így tovább folytathatják működésüket annak ellenére, hogy újra és újra pénzügyi nehézségekkel küzdenek. Szymanski ennek az elméletnek az érvényességét kérdőjelezi meg, angol futballklubok fizetéseképtelenségi eseteit és az ehhez kapcsolódó igazgatósági, illetve edzői elbocsátásokat elemezve. Eredményei azt mutatják, hogy a fizetéseképtelenséget követően számos igazgatót és edzőt elbocsátanak, vagyis a valóságban komolyabb személyi következményekkel jár a pénzügyi válság. A tanulmány egy alternatív magyarázatot is felvázol a klubok pénzügyi instabilitására. E szerint a futballklubok pénzügyi nehézségei gyakran az "előre nem látható" teljesítményalapú pénzügyi sokkokból erednek, különösen a kiesési és feljutási rendszerek miatt. Amikor egy klub egy alacsonyabb osztályba esik vissza, a bevételei drasztikusan csökkennek, míg a játékosokkal kapcsolatos költségek nem csökkennek azonnal. Ez különösen a kisebb, alacsonyabb osztályokban szereplő klubokra jelent nagy terhet, és hozzájárulhat a pénzügyi válságaikhoz (Szymanski, 2022).

Az UEFA jelentése alapján a 2023-as pénzügyi eredmények pozitív hírekkel szolgálnak, mivel a klubok fokozatosan visszatérnek a pénzügyi egyensúlyhoz a 2020, 2021 és 2022-es évek rendkívüli veszteségei után, amelyeket a pandémia okozott. A pandémia két fő módon érintette a klubokat: közvetlen bevételkiesések és folyamatos bérnövekedések miatt jelentős működési veszteségek, valamint a tranzfer tevékenység csökkenése miatt, ami átmeneti összeomlást eredményezett a tranzfer nyereségekben és bevételekben, tovább súlyosbítva az adózás előtti veszteségeket. A 2022-es évben a klubok 0.9 milliárd eurós működési veszteséget szenvedtek el a gyenge költségkontroll és a pandémia maradványhatásai miatt. A top három játékos-importáló ligák klubjai működési nyereséget jelentettek, míg a játékos-exportáló ligák klubjai általában működési veszteségeket könyveltek el. 2022-ben az angol és olasz klubok átlagosan több mint 50 millió eurós tranzferköltséget jelentettek, amelynek nagy része amortizációs költség volt.

A Premier League klubok pénzügyi helyzetének elemzése során megvizsgáljuk a klubok működési jövedelmezőségét, tranzfer tevékenységeit, valamint nem működési nyereségeit és veszteségeit, amelyek az adózás előtti nettó eredményhez vezetnek. A tranzfer tevékenység pénzügyi hatása összetett, mivel különböző módokon befolyásolja a pénzügyi eredményeket. Az egyszerűség kedvéért a tranzfer költségeket és bevételeket csoportosítottuk, amelyek nettó hatása közvetlenül befolyásolja az eredménykimutatást. Fontos megjegyezni, hogy a tranzfer tevékenység pénzügyi hatása jelentősen eltérhet a nyilvánosság által látottaktól, mivel a

költségek későbbi években történő fázisossága és az azonnali nyereség elismerése az értékesítés időpontjában más megvilágításba helyezik az adatokat.

Összességében a fejezet azt mutatja, hogy a klubok pénzügyi helyzete javuló tendenciát mutat 2023-ban, különösen a transzfer piac helyreállása és a költségek jobb kontrollja révén. Ez a tendencia azt sugallja, hogy a klubok sikeresen alkalmazkodtak a pandémia okozta pénzügyi kihívásokhoz, és erősebb pénzügyi alapokkal néznek szembe a jövőben.

2.2 A Premier League létrejötte

Az 1990-es évek elején az angol első osztályú klubok úgy vélték, hogy a labdarúgás radikális átalakítására van szükség ahhoz, hogy ők és általában a játék fejlődhessen és virágozhasson. Az első osztályú csapatok az 1991-92-es szezon után hozták létre a ligát azzal a céllal, hogy maximalizálják az angol labdarúgás gazdasági potenciálját. Az új liga hamarosan gazdag közvetítési és szponzori szerződéseket kötött, növelte a stadionok kényelmét és biztonságát, és a világ legjobb játékosai és menedzserei közül sokakat kezdett el vonzani. A Premier League minden egyes klubja a szezon során egy hazai és egy idegenbeli mérkőzést játszik a liga minden más csapata ellen. A győzelem három pontot jelent a győztes félnek a tabellán; döntetlen esetén minden klub egy pontot kap. A Premier League bajnoka az a csapat, amelyik a szezon végén a legtöbb ponttal rendelkezik. A liga legrosszabb három klubja kiesik, a másodosztályba (EFL Championship), míg a liga legjobb csapatai az UEFA versenysorozataiban indulhatnak el következő szezonban. (Premier League, 2024)

Ma már a Premier League részvénytársaságként működik. Az első televíziós megállapodást a fizetős televíziós műsorszolgáltatóval, a BSkyB-vel kötötték, ami akkoriban radikális lépésnek számított, de a Sky azóta is folyamatosan jelen van, és kulcsszerepet játszott a liga jelenlegi szintre való fejlesztésében. A Premier League megalakulásakor mindig is az volt a szándék, hogy a klubok számát 20-ra csökkentsék, hogy elősegítsék a fejlődést és a kiválóságot klub- és nemzetközi szinten. Ez az 1994/95-ös szezon végén valósult meg, amikor négy klub kiesett és csak kettő jutott fel. (Premier League, 2024)

Egy kis kitekintőként meg kell említeni, hogy hamarosan új változások jöhetnek a liga pénzügyi szabályozásában. Az új fizetési plafon szabályok a jelenlegi nyereség- és fenntarthatósági szabályok (PSR) helyébe lépnének. A javasolt szabályok közé tartozik, hogy az európai versenyeken nem induló klubok a bevételeik legfeljebb 85%-át költhetik a csapatok költségeire, beleértve a játékosok bérét, az átigazolási díjakat és az ügynöki díjakat, illetve az európai versenyeken részt vevő klubok bevételeik legfeljebb 70%-át költhetik ezekre a költségekre, az UEFA szabályainak megfelelően. A nagyobb klubok egyelőre a tervezet ellen

vannak, ezért Premier League vezetői finomítani fogják a javaslatokat, és a közelgő közgyűlésen igyekeznek jóváhagyni azokat. Ha jóváhagyják, az új szabályozás a 2025/26-os szezontól lép hatályba. (Solhekol, 2024)

2.3 Empirikus szakirodalom

2.3.1 A kapcsolat és aktualitása

Thakkar és Shakh (2021) tanulmánya szerint az adatok korának beköszöntével a vállalatok mára teljesen megváltoztatták működésük megközelítését, hogy maximalizálják nyereségüket és versenyelőnyre tegyenek szert riváisaikkal szemben. Ez természetesen igaz a sportszervezetekre is. Az adatvezérelt döntések alkalmazásával, amelyek az adattudomány, a Big Data és a gépi tanulás fogalmainak a döntéseikbe való beépítése eredményeként születtek, a sportcsapatok is leartadják a hasznukat. A sportolók jobbak, mint valaha, és ez nagyrészt az általuk elfogadott gyakorlatoknak köszönhető, amelyekben minden mozdulatot alaposan megvizsgálunk, nagy mennyiségű adatot generálnak, elemeznek, és ezek alapján fejlődnek (Thakkar és Shakh, 2021).

A labdarúgás első nagy adatelemzőjének egyértelműen Charles Reep nevezhető; ő volt az, aki legelőször próbált meg adatelemzést végezni annak érdekében, hogy a csapata játékát segíteni tudja. Reep többek között azt is vizsgálta, hogy hány passzal lehet a leghatékonyabban gólt szerezni egy szituációból, de még nem tett különbséget a passzok mérete, iránya és stílusa között, és emiatt következtetései sem voltak mindig helyesek, de mindenképpen úttörőnek számított a szakmában (Pollard, 2002). Reeptől a haladó statisztikákat felhasználó mai „laptop edzőkig” természetesen egy több mint 50 éves út vezetett, megannyi lépcsőfokkal és úttörő személlyel.

Az adatalapú klubmodell sikeres működésére lehetnek remek példák a Brighton és a Brentford, amely csapatok néhány év alatt eljutottak az alsóbb osztályokból a stabil Premier League felsőházi csapatok státuszába (a Brentford kicsit le van még ebben a folyamatban maradva, de közel vannak hozzá). A két klub tulajdonosai Tony Bloom és Matthew Benham egykor munkaviszonyban álltak, később pedig egymás riválisai lettek a sportfogadási piacon, majd pedig a futballban is. Korábbi céges és sportfogadási tapasztalataiknál már alkalmazták az adatalapú működést a saját cégjeiknél és ezt a tapasztalatukat kamatoztatják most a saját klubjaiknál is (Naylor és Harris, 2023).

Ezenkívül fontos még az elején letisztázni, hogy a labdarúgásban sportszakmai mutatókat megfigyelni, esetleg létrehozni nehezebb, mint sok más sportnál. Kevesebb benne az egyszerűbb és könnyen összehasonlítható jelenet más sportokhoz képest, ezért is van az, hogy

például a baseball ilyen téren sokkal előrébb tart, ott majdnem ugyanaz a jelent zajlik le egy meccsen több százszor. A foci egyrészt a várakozás sportja, másrészt pedig hatalmas benne a szerencse jelentősége; a véletlenek pedig mindig nagyon rontják a becslést (Szabó et. al, 2022). Bill Gerrard, a futballstatisztikák kutatója megjegyzi, hogy kihívást jelent a különböző játékosok cselekedeteinek súlyozása és a csapatok és ligák között alkalmazható univerzális mérőszámok elérése. A labdarúgás kiszámíthatatlan természete és a játékosok mozgásának kölcsönös függősége tovább bonyolítja a helyzetet (Phillips, 2011)

2.3.2 A modern adatelemzés a labdarúgásban

A nagy áttörés a statisztikák számára 1996-ban következett be, amikor az Opta Consulting elkezdte gyűjteni a Premier League meccsadatait. A cég vezetésének fő célja az volt, hogy saját márkát építsen a futballranglisták létrehozásával. A Premier League szponzora, a Carling fizetett az Opta Indexért, míg a klubok és a média ingyen kapta az adatokat. Minden klub kapott egy Excel-jelentést néhány alapvető statisztikával. Mai mércével mérve ezek a statisztikák kezdetlegesek voltak, de akkoriban ez forradalmi volt. Néhány év alatt a játék az adatok szűkösségéből a túl sok információ állapotába került. Hamarosan az Opta és riválisai hetente több ezer új adatpontot küldtek ki. (XFB Analytics, 2022)

Amennyiben az adatok típusait akarjuk csoportosítani a labdarúgásban, akkor 3 fő kategóriát lehet képezni: event data – labdás eseményalapú adatok, tracking data – pozíciós adat, illetve a fitneszadatok (Szabó et. al, 2022). Az első kategóriába eső adatok azok, amik a legtöbb szurkoló számára is elérhetőek a különböző statisztikai oldalakon. Ezen eseményadatok elemzése és feldolgozása összetettebb és időigényesebb, mint a tisztán mennyiségi statisztikai adatok elemzése. Ennek oka, hogy meccsenként több ezer eseményt és több százezer adatpontot rögzítenek. Minden egyes akciót, például passzt, cselezést, lövést, elkapást stb. rögzítenek, beleértve az x- és y-koordinátákat vagy a kezdő- és végkoordinátákat (pozíciót). Minden egyes akció esetében azt is rögzítik, hogy sikeres volt-e vagy sem.

Ezzel szemben az egészségügyi vagy fitneszadatok jellemzően csak a kluboknak állnak rendelkezésre, ezeket nem szokták publikálni. Ide tartoznak a különböző orvosi tesztek, vérvételek, de az úgynevezett „katapult” mellény is segít az adatgyűjtésben. A mellényben található szenzorok mérkőzések és edzések közben folyamatosan mutatják a játékosok legfőbb élettani adatait. Ezen számok alapján tudnak kalkulálni az edzői stábok, hogy ki mennyire terhelhető, illetve milyen állapotban, formában van. A katapult ezenfelül futási sebességet, mennyiséget és irányt is tud mérni (Catapult, 2024). Ezt pedig már a pozíciós adatoknál is felhasználják. Ennek egy alternatív mérési módja lehet a stadionok, edzőpályák felett

elhelyezett kamerák is (Szabó et. al, 2022). Ezek alapján határozhatóak meg az átlagos pozíciók, hő térképek stb.

A sport és adatelemzés közti kapcsolat ugyan már most jelentős fejlődésen ment keresztül, de még mindig van hova fejlődni. Szakértők a jövőt a multidiszciplináris megközelítésben rejlik, ahol a csapatok a sport és a tudomány összes sarokkövét egy közös tető alá tudják vonni fejlődnek (Thakkar és Shakh, 2021). Egy ilyen próbálkozást fedezhetünk fel az egyik angol top klubnál; a Liverpoolnál is. Jürgen Klopp edzősége alatt, már alaptól sokat fejlődött a klub adatelemző részlege, és gyakran használták fel az itt szerzett tudást a mérkőzéseken is, de ezzel még nincsenek megelégedve teljesen, és egy következő szintre akarják emelni az adatelemző részlegüket. A TacticAI nevű mesterséges intelligenciát is be szeretnék építeni a mindennapos taktikai felkészüléseikbe, ami jelen állapota szerint valószínűleg a rögzített szituációknál tudna nagy segítséget nyújtani (Tiwari, 2024). Bár a labdarúgó klubok egyre inkább elfogadják a statisztikai elemzést a teljesítmény javítása érdekében, a sportág eredendő összetettsége és a kereskedelmi érdekek jelentős kihívásokat jelentenek. A labdarúgás számszerűsítésére irányuló törekvés gyakran inkább titokzatossághoz vezet, mint tisztánlátáshoz, és a szurkolókat így elidegenítheti a szurkolókat a kedvenc csapataiktól (Phillips, 2011).

2.4 Teljesítményre ható mutatószámok a szakirodalomból

A dolgozat alábbi fejezetében összegyűjtöttem azokat a legfontosabb mérőszámokat, amelyek a szakirodalom szerint jelentős hatással vannak egy labdarúgócsapat teljesítményére. Ezeket a mutatókat az elemzés során is felhasználtam. Először a gazdasági oldalát mutatom be a teljesítményre történő hatásoknak, majd áttérek a sportszakmai témákra is.

2.4.1 Gazdasági hatás

Ami a legegyszerűbben elfogadott és alátámasztott pénzügyi befolyásoló tényező az adott kérdésben, az a játékosok fizetése. Ezzel szemben a játékosok átigazolására költött összegek meglepő módon nem egyértelműen jelentik azt, hogy egy csapat milyen teljesítményt fog nyújtani, a szakirodalomban erről eltérő vélemények vannak. Yaldo és Shamir (2017) tanulmánya mintafelismerő algoritmusokat használ a játékosok teljesítményadatainak elemzésére és a fizetések képességek alapján történő meghatározására, és erős korrelációt mutat a megjósolt és a tényleges fizetések között. Kiemeli, hogy a játékosok fizetése nemcsak a készségeket tükrözi, hanem a csapatteljesítményt is befolyásolja, hangsúlyozva a gazdasági mutatók fontosságát a futballcsapatok dinamikájának megértésében. Ferri és szerzőtársai (2017) szerint pozitív korreláció van a futballklubok játékosok fizetésére fordított kiadásai és a

játékosok sportteljesítménye között. A tanulmány azonban nem talált szignifikáns kapcsolatot a játékosok átigazolási díjai és a sportteljesítmény között. Ez arra utal, hogy a játékosok fizetéseibe való befektetésnek nagyobb hatása lehet egy labdarúgócsapat teljesítményére, mint az átigazolási díjakra fordított kiadásoknak.

Wanat és Leksowski (2022) az átigazolásoknak a futballklubok sportteljesítményére gyakorolt hatását tárgyalja, és összefüggést mutat ki a pénzügyi tevékenységek, például az átigazolások és a futballklubok sportteljesítménye között. Ez arra utal, hogy a gazdasági mutatók, különösen az átigazolásokhoz kapcsolódóan, jelentősen befolyásolhatják egy labdarúgócsapat teljesítményét. Prayoga és szerzőtársai (2022) rávilágítanak arra, hogy a sportteljesítmény és a pénzügyi mutatók jelentős hatására az európai futballklubok részvényárfolyamaira. A kutatás szerint a sportteljesítmény pozitívan befolyásolja a részvényárakat, míg a pénzügyi mutatók, például a likviditási és fizetőképességi mutatók szintén döntő szerepet játszanak. Úgy tűnik azonban, hogy a jövedelmezőségi mutatóknak nincs közvetlen hatásuk a részvényárfolyamokra. Ez aláhúzza mind a sport-, mind a gazdasági tényezők fontosságát a futbalcsapatok tőzsdei teljesítményének meghatározásában.

Keleş (2020) szerint a sportgazdaság, amely egyre nagyobb részesedést szerez a globális gazdaságból, várhatóan nagyobb előnyöket kínál a nagyobb költségvetéssel rendelkező profi labdarúgócsapatoknak. Ez arra utal, hogy az olyan gazdasági mutatók, mint a bevételek és a pénzügyi források szerepet játszanak. Szymanski és Kuypers (2014) szerint a csapatok bértömege és a nettó éves átigazolási kiadások olyan jelentős tényezők, amelyek pozitívan befolyásolják egy labdarúgócsapat sportsikerét (idézi Rohde és Breuer, 2016). Ez rávilágít az olyan gazdasági mutatóknak, mint a csapatok bérszámlái és az átigazolási kiadások a labdarúgócsapatok teljesítményére gyakorolt hatása szignifikáns.

Wakefield és Sloan (1995), valamint Matsuoka és szerzőtársai (2003) szerint a csapat teljesítményével való szurkolói elégedettség jelentősen befolyásolja a látogatottságot és más fogyasztási magatartást. A tanulmányok kiemelik továbbá a rajongók csapattal való azonosulásának fontosságát a hosszú távú kapcsolatok ápolásában és a vásárlók megtartásában a kapcsolati marketing stratégiák révén. Vagyis a csapat sikerességéhez nagyban hozzájárulhat egy lojális szurkolói bázis. Trail (2017) a sportszervezetek marketingjének területét vizsgálta, különös tekintettel a csapat identifikációjának és a csapat teljesítményével való elégedettségnek a különböző szurkolói magatartásokra, mint például a személyes látogatás, a média alapú látogatás, az áruvásárlás és a szájról-szájra történő kommunikáció hatására. Trail a szurkolói aktivitás és a csapatteljesítmény közötti bonyolult kapcsolatot mutatja be, és ő is hangsúlyozza,

hogy mind a helyi szurkolás, mind a globális jelenlét milyen pozitív hatásokat tud okozni egy klubnak.

Gómez és szerzőtársai (2018) szerint az átigazolási kiadások és a klubok teljesítménye közötti összefüggés jelentős, a magasabb átigazolási kiadásokkal rendelkező klubok általában jobb eredményeket érnek el mind a hazai, mind a nemzetközi versenyeken. A tanulmány arra is utal, hogy az átigazolási piaci stratégiák jelentős szerepet játszanak a klubok teljesítményének alakításában, ami a klubok és a ligák közötti egyenlőtlenségekhez vezethet, és mint utóbb az élet is igazolta, vezetett is. Liu és szerzőtársai (2016) szerint a modern profi labdarúgás „pénzjáték”, amelyben a tehetséges játékosok megszerzésére fordított nagyobb befektetések általában jobb csapatteljesítményt eredményeznek. Azonban azt is kimutatták, hogy a profi futballklubok valójában különböző stratégiákat játszhatnak a túlélés vagy akár a kiemelkedő teljesítmény érdekében és a stratégiák sikere erősen összefügg a futballisták átigazolási hálózatában lévő hálózati tulajdonságaikkal.

A transzferdíjakra kitérve, a bevezetőben emlegetett „egy drágább játékos biztosan jobb játékos is, mint egy olcsóbb” feltevés több szempontból is téves feltevés tud lenni. Egy játékos értékét nem csupán a játékos minősége adja. Olyan tényezők is jelentősen befolyásolják, mint a játékos életkora, korábbi teljesítménye, lejátszott játékkercei, nemzetisége, válogatottsága, korábbi klubcsapatai stb. (Užik és munkatársai, 2022 és Franceschi és munkatársai, 2024). Ezenfelül azt is fontos megjegyezni, hogy az igazolásoknak közel sem válik be a száz százaléka. Dr. Ian Graham a Liverpool Fc egykori kutatási igazgatója szerint historikusan a bevált igazolások aránya nagyjából 50 százalékra tehető a 10 millió font feletti igazolásoknál (Hobbs, 2021) – a beváltságot egy saját felállított szempont rendszer alapján értékelte a szakértő. Hét okot sorolt fel, ami miatt megbukhat egy igazolás: 1. A jelenlegi játékos jobb, mint az új játékos 2. A játékos nem olyan jó, mint azt először gondolták 3. A játékos nem illik a csapat stílusához 4. A játékos nem a megfelelő pozícióban játszik 5. A menedzser nem értékeli a játékost 6. A játékosnak egészségügyi problémái vannak 7. A játékosnak személyes problémái vannak.

2.4.2 Sportszakmai hatás

Egy meccs attól tud izgalmassá és érdekessé válni, ha sok lövést, helyzetet és gólt láthatunk rajta. Ezenfelül még vannak más dolgok is, amiktől szépnek titulálhatjuk a játékot, például a sok passzos labdakihozatalok, nagy sprintek, védések, szerelések stb. Ezek viszont nem jelentik egyértelműen azt, hogy ha valamelyik esemény száma növekszik, akkor egyértelműen jobb lesz egy csapat. Ha például több becsúszószerelést kell egy csapatnak bemutatnia, az inkább arról árulkodik, hogy sokat támadtak ellenük, többet kellett védekezni.

Zambom-Ferraresi és szerzőtársai (2018) kutatásukban több fontos meghatározó indikátort is kiemeltek, ami befolyásolja egy csapat sikerességét, de kiemelték azt, hogy a támadó mutatók sokkal jelentősebbek, mint a védekezési mutatók (ez részben azért is van, mert a védekezést igazán jól leíró statisztikai mutatók jelenleg még nincsenek a sportban), így a kutatásomban én is jobban fogok az előbbire fókuszálni.

A szakirodalomban Collet (2013) például kiemelte, hogy fontos összefüggés van a csapatsiker és a passzolás között, illetve a több labdabirtoklás és a passzolási pontosság között is. Ezt Göral (2015) is alátámasztotta a 2014-es világbajnokság adataival. Itt figyelembe kell viszont venni, hogy egy világbajnokságon jóval kevesebb meccset játszanak a csapatok, mint egy egész szezonos bajnokságban, vagyis itt kisebb volt a minta. Ehhez hasonló kutatás volt Castellano és szerzőtársai (2012) kutatása; szerintük a labdabirtoklás a labdarúgásban kulcsfontosságú teljesítménymutató, amely különbséget tud tenni a győztes, a döntetlen és a vesztes csapatok között. A tanulmány három világbajnoki torna adatait elemezte, és megállapította, hogy a labdabirtoklás, valamint más támadójáték-változók, például az összes lövés és a kapura lövések száma döntő fontosságú a csapatok eredményeinek megkülönböztetésében. Ezzel aláhúзва a labdabirtoklás jelentőségét a csapatok teljesítményének elemzésében, és kiemeli a labdarúgócsapatok sikerének meghatározásában betöltött szerepét. Clemente és Martins, (2017) tanulmányukban azt találták, hogy a nagyobb labdabirtoklással rendelkező csapatok hajlamosabbak hosszabb passzsorozatokra, és több játékost homogénebb módon vonnak be, ami hatással lehet a csapat teljesítményére. A labdát többet birtokló csapat könnyebben tud eljutni az ellenfél kapujáig és így gólt szerezni, illetve amíg náluk van a labda, addig gólt sem tudnak kapni, így elérve meccseken egy jobb eredményt.

Lepschy és szerzőtársai (2021) a 2014-es és 2018-as világbajnokságokat vizsgálták. Összesen 128 mérkőzést elemeztek általánosított sorrendű logit módszerrel. Huszonkilenc változót azonosítottak a korábbi kutatásokból. Az eredmények azt mutatták, hogy a védelmi hibák, a gólhatékonyság, a párharcok sikeressége, a szerelések sikeressége, az ellentámadásokból származó lövések, a tisztázások és a keresztpasszok jelentős hatással voltak a mérkőzés megnyerésére az említett tornák során. A labdabirtoklás, a távolság és a csapatok piaci értéke nem volt hatással a sikerre. Összességében a legtöbb kritikus sikertényező és a szoros mérkőzések megnyerésére legnagyobb hatást gyakorló tényezők a védekező akciók voltak. Az eredmények továbbá arra utalnak, hogy a közvetlen játék és a presszing hatékonyabb volt, mint a labdabirtoklásos játék.

Modric és szerzőtársai (2021) szerint bár a futóteljesítmény és a labdabirtoklás, valamint a csapatteljesítmény közötti összefüggéseket a labdarúgásban gyakran feltételezik, az elit

labdarúgásban az ezzel kapcsolatos tényleges ismeretek korlátozottak, ezért ők ezt vizsgálták az UEFA Bajnokok Ligájában. Az eredményeik azt mutatták, hogy a teljes, a gyaloglás és a kocogás során megtett távolságok negatívan és pozitívan kapcsolódtak a csapatteljesítményhez. A megnyert mérkőzéseket a teljes, a gyaloglás és a kocogás során megtett távolságok szignifikánsan alacsonyabb, illetve magasabb értékei jellemezték.

A szakirodalom számos mutatót alkalmaz a klubok játékoskeretének értékelésére, köztük a játékosok átlagéletkorát. Ennek a mutatónak a csapatteljesítménnyel való kapcsolata nem egyértelmű, mivel Dendir (2016) kutatása szerint a profi labdarúgók általában 25 és 27 éves koruk között vannak karrierjük csúcsán, ami hozzávetőlegesen pályafutásuk közepére esik. Ez az életkori csúcs azonban egyénenként változhat, mivel számos külső tényező – például súlyosabb sérülések – módosíthatja a karrier alakulását. Torgler és Schmidt (2013) szerint az egyéni teljesítmény és a csapattársak életkora között statisztikailag szignifikáns korreláció található, ami arra utal, hogy a keret átlagéletkora a csapatsikerre is hatással lehet. Az összefüggés iránya akár azt is megmutathatja, hogy a fiatalabb vagy az idősebb játékoskerettel rendelkező klubok sikereesebbek-e.

Fontos indikátor, amiről nem esett még eddig szó a lövések és kapura lövések száma. „Lövésből lesz a gól” szól a szállóige, melynek igazságtartalma vitathatatlan, de azért ennél egy fokkal bonyolultabb a helyzet. Ugyan Rocha-Lima és szerzőtársai 2021-es kutatásukban, illetve Zambom-Ferraresi és szerzőtársai 2018-as elemzésénél azt az eredményt kapták a kutatók, hogy a kapura tartó lövések száma befolyásolja egy csapat sikerességét, ezek a kutatások még nem vizsgálták az egyes lövések minőségét azontúl, hogy megkülönböztettek kaput eltaláló és nem eltaláló lövéseket (érdemes megjegyezni, hogy például a kapufák nem számítanak kaput találó lövéseknek – OPTA, 2024). Az anekdota szerint anno Charles-Reep sem tett még ilyen szinten különbséget a lövések között, pedig a modern sportszakmai irányok szerint ebben lehet a kulcs. Az újfajta haladó statisztikai mutatók, mint például az expected goal modell (xG) alapján változtatta meg az elmúlt évek labdarúgását. (Analyst, 2024)

Az xG egy olyan mutató, amely megbecsüli a lövések valószínűségét, hogy várhatóan mekkora eséllyel azokból lesz gól (Whitmore, 2023). Minden lövéshez egy xG-értéket rendelnek, amely 0 és 1 között mozog; az érték annál magasabb, minél nagyobb eséllyel születik belőle gól. Az xG-t különféle tényezők alapján számítják ki, mint például a lövés helyzete, a támadó szöge, az ellenfél kapusának pozíciója és a védők elhelyezkedése stb. A különböző adatelemző cégek és oldalak (például: Opta, Statsbomb, Infogol) különböző modelleket – más paramétereket használnak, melyek értékei nagyságrendileg megegyeznek egymással, de minimális eltérés azért megfigyelhető; jó példa erre a büntetőrúgásért adott xG érték, amely

jellemzően 0,74 és 0,8 érték között mozog a különböző elemzői oldalakon (Szabó et. al, 2022). Tiippana (2020) a különböző lövésekkel és gólhelyzeteket használó modelleket hasonlított össze; köztük az xG és egy egyszerű lövéseken alapuló modellt is. Az eredményei szerint az xG modell kifejezetten a teljes szezonos becslés esetében bizonyult pontos modellnek. A további haladó statisztikai mutatót az adatbázisom bemutatásánál tárgyalom, mivel ezen mutatók alkalmazása a szakirodalomban, még nem mondható elterjedtnek.

Szakirodalmi áttekintőm lezárásához közeledve, és a módszertan fejezetre ráfordulva még röviden összefoglalnám, hogy korábbi más kutatásokban az általam is használt módszertant; a panel ökonometriai elemzést mire használták fel és milyen eredmények születtek. Gásquez és Royuela (2016) tanulmánya a nemzetközi futball sikerének gazdasági és társadalmi meghatározóit vizsgálta az ELO-rangsor adataival. A szerzők panelökonómiai kutatást alkalmaznak, hogy elemezzék, a gazdasági fejlettség és társadalmi tényezők milyen mértékben befolyásolják az országok futballteljesítményét. Eredményeik szerint a magasabb GDP-vel és egy főre jutó jövedelemmel rendelkező országok általában előkelőbb helyezéseket érnek el, a tanulmány ezzel azt hangsúlyozza, hogy a nemzeti gazdasági jólét összefügg a nemzetközi sportban elért sikerrel. Egy másik nemzetközi területet vizsgáló tanulmányban Valenti és szerzőtársai (2020) női nemzeti labdarúgó-válogatottak teljesítményét vizsgálták, a sportpolitikák és a nemzetközi siker kapcsolatára összpontosítva. Panel adatok segítségével elemzik, hogy az elit sport támogatása, valamint az infrastrukturális és anyagi befektetések milyen hatással vannak a válogatottak eredményességére. Az eredmények alapján azok az országok, amelyek jelentős anyagi támogatást és fejlett sportinfrastruktúrát biztosítanak, jobb eredményeket érnek el a női futballban.

Az előbb említett szakirodalmak másfajta alkalmazásokat, területeket mutattak be, de vannak, amelyek szorosabban kapcsolódnak a dolgozatomhoz. Ilyen Beck és Meyer (2012) kutatása, ahol a futballcsapatok teljesítményének modellezését tűzték ki célul, elemzésükben elméleti és empirikus szempontokat kombinálnak, hogy bemutassák, hogyan lehet a csapatok eredményességét kvantitatív módszerekkel mérni és előre jelezni. A tanulmány rávilágít arra, hogy a paneladatok felhasználása sokkal pontosabb teljesítményértékelést tesz lehetővé a csapatsportokban, mint az egyedi megfigyelések. Findikçi és Tapşın (2015) ugyancsak panelökonómiai módszereket alkalmazva vizsgálták, hogy a gazdasági tényezők – mint a klub költségvetése, játékosfizetések és befektetések – hogyan kapcsolódnak a csapatteljesítményhez. A kutatás eredményeik ugyanazt mutatják, mint amit a fennebbi fejezetben bemutattam; a magasabb költségvetéssel rendelkező csapatok hajlamosak jobb eredményeket elérni, és ez alátámasztja azt az elképzelést, hogy a gazdasági erőforrások jelentős szerepet játszanak a

sportsikerek elérésében. Ezek a kutatások is megmutatják, hogy a panelökonómiai elemzés jól alkalmazható különféle szintű és jellegű labdarúgó-teljesítmények és a siker befolyásoló tényezőinek azonosítására, legyen szó nemzetközi vagy válogatott szintről, férfi vagy női csapatokról, illetve különböző gazdasági és sportpolitikai összefüggésekről.

3. Módszertan

A kutatási kérdésem az volt, hogy melyek azok a mutatók a modern labdarúgásban vizsgált gazdasági és sportszakmai statisztikai mutatók közül, amelyek szoros kapcsolatban állnak egy csapat teljesítményével. Ennek megválaszolására egy olyan modellt állítok fel, ahol az eredményváltozóm a szezononként elért pontszám lesz, magyarázó változóimat, pedig a szakirodalomban is bemutatott mutatókból és más gazdasági és sportszakmai változókból állítom össze. Az adatgyűjtést követően az elemzésem során panel ökonometria modellezést fogok használni, hogy a lehető legpontosabb eredményeket kapjam. A panel adatok olyan megfigyeléseket tartalmaznak, amelyek több egységre (pl. személyekre, vállalatokra vagy országokra) vonatkoznak egy adott időintervallumban, lehetővé téve a változók időbeli és keresztmetszeti különbségeinek alapos elemzését. A módszertan fejezetben Wooldridge (2013) és Griffiths és munkatársai (2012) által írt tankönyveket használom fel. A kutatást az R Studio programmal végeztem, az adatok nagy részét a worldfootballR nevű R csomaggal gyűjtöttem össze (Zivkovic, 2022). A panel ökonometria vizsgálat előtt még leíró statisztikai elemzést fogok végezni, hogy amennyiben szükséges egyes változókat átalakítani, például logaritmálni, akkor az megtörténjen.

A panel modellezés alkalmazása indokolt a vizsgálatban, mivel az adatbázisomban egyszerre jelenik meg a keresztmetszeti és idősoros dimenzió, ami több információt nyújt, mint a kizárólag keresztmetszeti vagy csupán idősoros adatbázisok. A panel adatok segítségével a függő változóként szereplő elért pontszámot olyan magyarázó változók alapján becsüljük, amelyekről mind keresztmetszeti, mind idősoros információk rendelkezésre állnak. Emellett ebben a modelltípusban is vannak nem megfigyelt változók, amelyeket a hibatag foglal magában. A keresztmetszeti dimenzió csoportjai nálam a csapatok lesznek – viszont a kieséses rendszer miatt egyből lehet tudni, hogy a vizsgált egyének nem lesznek ugyanazok a teljes mintán, emiatt én unbalanced (kiegyensúlyozatlan) panel elemzést fogok végezni.

A panel adatok elemzéséhez leggyakrabban használt modellek közé tartozik a rögzített hatások (FE), véletlen hatások (RE) és az összevont legkisebb négyzetek (Pooled OLS) modell. Mindegyik modell eltérő feltételezéseket tesz a nem megfigyelt egyéni hatások és a magyarázó változók kapcsolatáról. A Fixed Effects modell lehetővé teszi, hogy minden egyes egységnek

külön-külön metszéspontja legyen, így megragadva az időben állandó egyéni különbségeket. Különösen hasznos, ha feltételezhető, hogy a nem megfigyelt egyéni jellemzők korrelálnak a magyarázó változókkal. A Random Effects modell az egyéni hatást egy véletlen változóként kezeli, amely független a magyarázó változóktól. Így lehetséges lesz a hatékony becslés általánosított legkisebb négyzetek módszerével (GLS). Azonban, ha a nem megfigyelt hatások korrelálnak a magyarázó változókkal, az RE becslés inkonzisztens lesz, így a FE vagy az összesített OLS alkalmazása előnyösebb. A Pooled OLS használata akkor lehet indokolt, ha egyik modell feltételei sem teljesülnek, vagy ha az egyéni hatások jelentéktelenek. Ez az eljárás minden megfigyelést összegyűjt, figyelmen kívül hagyva az egyéni hatásokat. Ugyanakkor, ha az egyéni hatások és a magyarázó változók között korreláció áll fenn, az eredmények torzítottak lehetnek.

Az alábbiakban a modellek általános egyenleteit mutatom be. Az összevont (Pooled OLS) modell általános egyenlete:

$$y_{it} = \beta_0 + \beta_1 x_{it1} + \beta_2 x_{it2} + \dots + \beta_k x_{itk} + u_{it}$$

Itt y_{it} az i -edik egység (pl. csapat) t -edik időszakra vonatkozó függő változója, x_{itj} pedig a magyarázó változók. Ez a modell feltételezi, hogy minden egyén és időszak esetében az együtthatók azonosak. Ehhez képest a rögzített hatású (FE) modell egyenlete:

$$y_{it} = \alpha_i + \beta_1 x_{it1} + \beta_2 x_{it2} + \dots + \beta_k x_{itk} + u_{it}$$

A különbség az előző egyenlethez képest a β_0 és az α_i cseréje. A rögzített hatások modellben az α_i egyéni hatások, amelyek minden egységhez külön metszéspontot rendelnek, így figyelembe véve az időben állandó, de megfigyelhetetlen egyéni jellemzőket. Ez a modell különösen akkor előnyös, ha az egyéni hatások korrelálnak a magyarázó változókkal. A véletlen hatású (RE) modell gyakorlatilag ötvözi a két fenti egyenletet; itt egyszerre lesz β_0 és α_i is a modellegyenletben:

$$y_{it} = \beta_0 + \beta_1 x_{it1} + \beta_2 x_{it2} + \dots + \beta_k x_{itk} + \alpha_i + u_{it}$$

A véletlen hatású modellben az α_i véletlen változóként szerepel, és független a magyarázó változóktól. Ha az α_i és a magyarázó változók között nincs korreláció, akkor a RE modell hatékonyabb becslést biztosít a generalizált legkisebb négyzetek (GLS) módszerével.

Arról, hogy mely modellt kell használnunk 3 teszttel tudunk döntést hozni; Chow, Hausman és LM tesztekkel. Először a Chow-teszt alkalmazásával dönthetünk arról, hogy a Pooled OLS vagy a fix hatású (FE) modell-e a megfelelő. A teszt nullhipotézise szerint minden egyénre és minden időszakra azonosak a magyarázó változók együtthatói, vagyis a meredekség

azonos az összes csoport esetében: $H_0: \beta_i, k = \beta_k$ minden egyénre és minden k magyarázó változóra. Ha a teszt eredménye $p > 0,05$, akkor a Pooled OLS modellt választjuk, mivel elfogadjuk a nullhipotézist. Ha viszont $p < 0,05$, akkor elutasítjuk a nullhipotézist, és a fix hatású (FE) modellt választjuk, mert ebben az esetben az egyéni hatások figyelembevétele szükséges.

Ha a Chow-teszt alapján fix hatású modell tűnik helyesnek, akkor a Hausman-teszt következik, amely eldönti, hogy a fix hatású (FE) vagy a véletlen hatású (RE) modell a megfelelő választás. A Hausman-teszt nullhipotézise azt feltételezi, hogy az egyéni hatások nem korrelálnak a magyarázó változókkal. Ha a teszt eredménye $p > 0,05$, akkor az RE modellt választjuk, mivel az egyéni hatások és a magyarázó változók között nincs korreláció, így az RE modell hatékonyabb. Amennyiben elutasítjuk a nullhipotézist, akkor a fix hatású (FE) modellt választjuk, mert a magyarázó változók és az egyéni hatások között korreláció áll fenn.

Végül, ha a véletlen hatású modell és a Pooled OLS között kell dönteni, a Lagrange Multiplier (LM) teszt alkalmazható. Az LM-teszt nullhipotézise szerint nincs szignifikáns egyéni hatás, azaz a Pooled OLS a megfelelő modell. Ha elfogadjuk a H_0 -t, akkor a Pooled OLS modellt választjuk, mert nem szükséges figyelembe venni az egyéni hatásokat. Ezzel szemben, ha elutasítjuk a nullhipotézist, akkor a véletlen hatású (RE) modell mellett döntünk, mivel a szignifikáns egyéni hatások jelenléte indokoltta teszi a RE modell alkalmazását.

A modellekben található heteroszkedaszticitás tesztelésére a Breusch-Pagan tesztet fogom alkalmazni, amelynek nullhipotézise szerint a hibatagok varianciája állandó. Ez azt jelenti, hogy az adatok hibatagjainak varianciája viszonylag állandónak tekinthető, és a modell becslései a klasszikus feltételezések szerint érvényesek maradhatnak. Amennyiben szükséges a robusztusabb ellenőrzés érdekében HC1 és HC3 típusú robusztus standard hibákat is alkalmazni fogom, hogy biztosítsam az eredmények stabilitását még akkor is, ha rejtett heteroszkedaszticitás lenne jelen. Ezzel a módszerrel minimalizálhatók a standard hibák torzításaiból adódó problémák, így a modellünk eredményei megbízhatóbbá válnak.

A jelen kutatásomban különféle mértékegységű változókat használok (pl. gólok, százalékok, piaci értékek), így a standardizálás fontos ahhoz, hogy a modell eredményei ne torzuljanak. Az irodalmi források áttekintése alapján ez az elemzési szempont a meglévő kutatásokban eddig nem kapott kiemelt figyelmet, ezért a standardizálás alkalmazása új aspektust jelenthet, amely segíti az elemzés során a változók relatív fontosságának pontosabb megértését, és lehetőséget biztosít arra, hogy az elemzés következő két modelljében is egységes összehasonlítást végezzek a kapott eredmények között.

A változók standardizálása széles körben használt módszer a statisztikai elemzésekben, különösen, ha eltérő skálájú vagy mértékegységű adatokat kezelünk. A standardizálás során

minden változót azonos skálára helyezzünk, azaz az átlagukat 0-ra, szórásukat pedig 1-re állítjuk úgy, hogy minden értékből kivonjuk a változó átlagát, majd elosztjuk a szórásával. Ezáltal a különböző típusú változók közvetlenül összehasonlíthatók, és a regressziós modellek koefficiensei is könnyebben értelmezhetők. A legnagyobb abszolút értékű standardizált koefficienssel rendelkező változó mutatja meg a legnagyobb hatást a függő változóra, mivel az jelzi, hogy a független változó egységnyi szórásnyi növekedése hogyan módosítja a függő változót.

4. Adatbázis bemutatása

Kutatásom célja, hogy meghatározzam azokat a legfontosabb indikátorokat, amelyek befolyással vannak a labdarúgó csapatok teljesítményére, ehhez pedig egyszerre használok fel sportszakmai és gazdasági mutatókat is. Adatbázisom 7 szezon adatait tartalmazza (2017/18-as szezontól kezdve a 2023/24-es szezonig) az összes Premier League-ben szereplő csapatra. A kutatásom alapjául 3 internetes adatbázis szolgált; a sportszakmai mutatókat az Fbref-ről (2024) gyűjtöttem össze, mely az OPTA adatait használja, a pénzügyi, gazdasági adatokat pedig a Transfermarkt (2024) és a Capology (2024) oldaláról – előbbiről az átigazolásokkal és játékosértékekkel kapcsolatos adatok származnak, utóbbiról a játékosok fizetési kerete. Az adatok nagy részét a worldfootballR (Zivkovic, 2022) nevű R csomaggal gyűjtöttem össze a kiválasztott időszakra, de néhány adatot kézi begyűjtéssel szedtem össze az oldalakról. Mivel minden vizsgált szezonban 38 meccset játszott le minden csapat, és a kutatásomban nem teszek különbséget a hazai és az idegenbeli pályán lejátszott mérkőzések között, ezért itt lehet egybe vizsgálni a teljes szezonos adatokat, nem kell leosztani meccsenkénti, azaz 90 perces mutatókra. Az adatbázisomban 31 változó szerepel, amelyek közül a többség a szakirodalomban előforduló mutatókból származik. Az elemzés szempontjából kiemelkedően fontos változó az elért pontszámok, ugyanis ez lesz az eredményváltozó, tehát ennek alakulását magyarázza majd a felépített modell a későbbiekben. A csapatok nevei, illetve helyezések az adott szezonban szintén fel van tüntetve a modellben, azonban ezen változók csak nominális változók, a helyezés ordinális, a többi változó pedig numerikus típusú. A változókat és a velük kapcsolatos várakozásaimat, hogy milyen hatással lesznek az eredményváltozóra az alábbi táblázatban foglaltam össze:

Rövidítés	Jelentés	Várakozás
<i>Season_End_Year</i>	A szezon befejezésének éve	
<i>Squad</i>	Csapat neve	
<i>Rk</i>	Helyezés (sorszám)	

<i>GF</i>	Lőtt gól (Goals For) (db)	+
<i>GA</i>	Kapott gól (Goals Against) (db)	-
<i>GD</i>	Gólkülönbség (Goal Difference) (db)	+
<i>Pts</i>	Pontok (Points) (db)	eredmény
<i>xG</i>	Várható gólok (Expected Goals)	+
<i>xGA</i>	Várható kapott gólok (Expected Goals Against)	-
<i>xGD</i>	Várható gólkülönbség (Expected Goal Difference)	+
<i>Attendance</i>	Látogatottság (db)	+
<i>Att_Total</i>	Összes látogatottság (db)	+
<i>Cmp_percent_Total</i>	Teljesített passzok százaléka (Completion Percentage Total)	+
<i>xA_Expected</i>	Várható asszisztok (Expected Assists)	+
<i>Sh_Standard</i>	Standard lövések (db)	+
<i>SoT_percent_Standard</i>	Standard kaput eltaláló lövések százaléka (Shots on Target)	+
<i>npxG_Expected</i>	Nem büntetőből származó várható gólok (Non-Penalty xG)	+
<i>Weekly_Wages£</i>	Heti bér Fontban	+
<i>Age</i>	Kor	?
<i>Poss</i>	Labdabirtoklás (Possession) (%)	+
<i>PrgC_Progression</i>	Labdacipelés (Progressive Carries) (db)	+
<i>PrgP_Progression</i>	Progresszív passzok (Progressive Passes) (db)	+
<i>SCA</i>	Várható gólhelyzetek (Shot-Creating Actions) (db)	+
<i>PPDA</i>	Passes Per Defensive Action - letámadás inetnizását mérő mutató	-
<i>Expenditure</i>	Kiadások (millió €)	+
<i>Income</i>	Bevétel (m €)	?
<i>Balance</i>	Egyenleg (m €)	-
<i>Squadsize</i>	Csapatméret (db fő)	?
<i>Foreigners</i>	Külföldi játékosok (Foreign Players) (db fő)	?
<i>ø market value</i>	Átlagos piaci érték (Average Market Value) (m €)	+
<i>Total market value</i>	Teljes piaci érték (m €)	+

1. táblázat: Változók bemutatása - Forrás: Saját szerkesztés

Szakirodalmi kutatásom során jellemzően csak egyszerűbb mutatókat használtak (labdabirtoklás, kapura lövés, passzpontosság, stb), viszont a modern labdarúgó adatelemzők trendjeihez alkalmazkodva én a vizsgálatomba már bevonom az újfajta haladó statisztikákat is, melyeket az OPTA leírásai alapján mutatok be (2024). A már fentebb bemutatott xG egy specifikus változata az npxG, azaz a non penalty xG, mely nem veszi figyelembe a büntetőrúgásokat; vagyis a „tisztá” támadásokat és helyzeteket veszi figyelembe (az idézőjel azért került be, mert más pontrúgásból jövő helyzeteket nehéz lenne tiszta támadásnak nevezni, de ettől most tekintsünk el). Ha a még tovább fejlesztett xG modellt vizsgáljuk eljutunk az xGOT, azaz expected goal on target mutatóhoz (másnéven PSxG – pre-shot xG), mely az összes lövést figyelembe véve kerül kiszámításra a lövés időpontjában, a lövési kísérlet minőségének ismerete nélkül. Nemcsak a célba talált lövéseket tartalmazza, hanem azon lövéseket is, amelyek „elhajlottak” vagy célt tévesztettek.

Ha már a gólokról beszéltünk, akkor meg kell említenünk az xA mutatót is, mely a várható gólpassz mutató - annak a valószínűsége, hogy egy adott befejezett passzból gólhelyzetet eredményező assziszt lesz. Ez az Opta által kifejlesztett statisztika minden passzhoz hozzárendel egy valószínűséget a passz típusa, a pályán való elhelyezkedés, a játék fázisa és a megtett távolság alapján. A játékosok minden befejezett passzért xA-t kapnak, függetlenül attól, hogy történt-e lövés vagy sem. Annak érdekében, hogy csak a lövéshez asszisztáló passzok xG-jét különítsék el, létezik az Expected Assisted Goals (xAG). Ez azt jelzi, hogy egy játékos mennyire képes gólszerzési lehetőségeket előkészíteni anélkül, hogy a lövés tényleges eredményére vagy a lövő szerencséjére/képességére kellene hagyatkoznia. A játékosok csak akkor kapnak xAG-t, ha egy befejezett passz után lőnek.

Az SCA (Shot-Creating Actions) egy olyan statisztikai mutató, amely azt méri, hogy egy játékos hányszor járult hozzá közvetlenül egy lövési helyzet kialakításához a mérkőzés során. Az SCA két olyan tevékenység – például passz, csel, szerelés vagy szabálytalanság kiharcolása – összekapcsolódását jelzi, amelyek közvetlenül lövési lehetőséghez vezetnek. Ezen mutató használata segíti annak megértését, hogy a játékosok hogyan vesznek részt a támadásokban, és hogyan teremtenek helyzeteket, függetlenül attól, hogy maguk lőnek-e kapura vagy sem. További támadást leíró mutatóként használok még a progresszív labdacipelést (Progressive Carries) és a progresszív passzokat (Progressive Passes).

A progresszív labdacipelés a labdarúgásban az a statisztikai mutató, amely azt méri, hogy egy játékos hányszor viszi előre a labdát jelentős távolságon (egészen pontosan öt méternél feljebb) keresztül, ezzel elősegítve a csapat támadását és a támadó zónák elérését. A

progresszív passzok, pedig olyan passzok, amelyek jelentősen előre viszik a labdát a pályán, közelebb az ellenfél kapujához.

Kutatásomban a védekezésre fókuszáló mutatók az xGA és PPDA – expected Goals Against, Passes Allowed per Defensive Action – vagyis a várható kapott gólok és az engedélyezett passzok száma védekező akciónként lesznek. A PPDA az ellenfél saját védekező harmadán kívül engedélyezett ellenfél passzok száma osztva az ellenfél saját védekező harmadán kívül végrehajtott védekező akcióinak számával. A PPDA-számításban a védekező akciók a szabálytalanságok, a szerelések, az elkapások, a kihívások és a blokkolt passzok szerepelnek (Analyst, 2024).

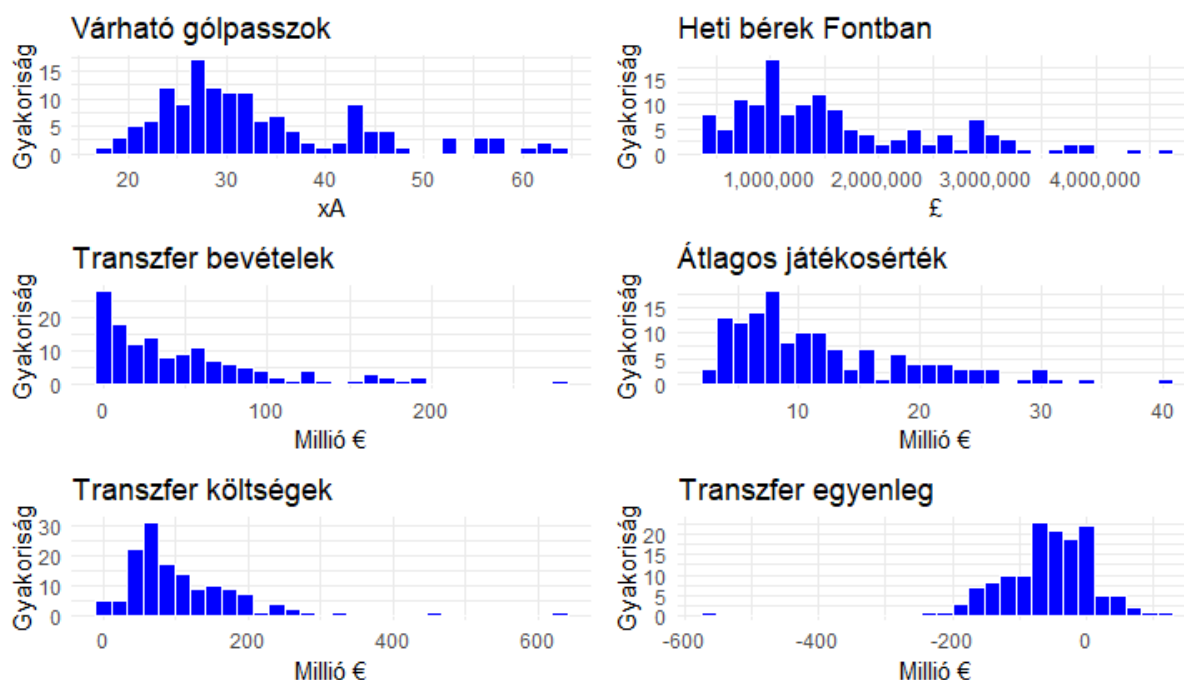
A várakozásnál az az alapfeltevés, hogy meccset úgy lehet könnyebben nyerni, ha támadóbb játékot játszik egy csapat, vagyis jobban törekszik a progresszivitásra. Mindazonáltal néhány mutatóval kapcsolatos várakozásom elsőre kérdéses vagy meglepő lehet. Példának okáért a kor azért lett kérdéses, mert a szakirodalomban is bemutatott szöveg alapján a játékosoknak nem lineárisan nő a minőségük az életkorukkal, ezért nem lehet eldönteni előre, hogy az milyen hatással lesz az eredményváltozóra, hasonló a helyzet a játékoskeret méreteire vonatkozó mutatókkal. A mérleget azért raktam negatívra, mert az a feltevés, hogy ha valaki nagyon pluszban zár egy átigazolási időszakot, akkor az azt jelenti inkább, hogy a csapatából csak elmentek a minőségi játékosok és nem sikerült őket pótolni az értük kapott pénzből.

5. Elemzés

5.1 Leíró statisztika

Kutatásom első lépéseként megvizsgáltam az adataimat a leíró statisztika szempontjából. Elsősorban outlier értékek, illetve aszimmetrikus eloszlások szűrésére használtam ezt a módszert. Az összes változónál - ahol lehetett és volt értelme – megnéztem az alap statisztikai mutatókat; ezek eredményeit a dolgozat végén a mellékletek fejezet tartalmazza. Az eredmények azt mutatják, hogy nem volt olyan outlier érték, amit érdemes lett volna kiszűrni, viszont néhány változó eloszlása jelentősen balra ferde, azaz jobbra elnyúló volt – ezt a ferdeség mutató alapján lehet következtetni, ha 1-nél nagyobb az érték akkor balra ferde az eloszlás, ha -1-nél kisebb, akkor jobbra ferde. Ahogyan az a legtöbb kutatás során elő szokott fordulni a pénzügyi mutatók esetében jelentkezett az aszimmetria; változóim közül a heti bérek, a transfer bevételek és kiadások, az átlagos piaci értékek és a várható asszisztoknál beszélhetünk egyértelműen balra ferde eloszlásokról, a transferegyenleg esetében, pedig jobbra ferde eloszlásról. A transferegyenleg ferdesége egy outlier értékkel magyarázható, amely a Chelsea 2022-2023-as szezonbeli költsége (a klub ekkoriban érkezett új tulajdonosa

elképesztően magas tőkét fektetett be a klubba). Az előbb kiemelt eloszlásokat az alábbi hisztogramok mutatják be:



3. ábra: Folytonos változók hisztogram ábrái - Forrás: Saját szerkesztés

A balra ferde eloszlások normalizálására az általános módszer az, hogy logaritmáljuk az adatokat, ezzel közelítve őket a normális eloszláshoz. A logaritmált változatokat már fel lehet használni a modellezéshez. A transzfer egyenleg változónál az outlier értéket meghagytam, mert túl nagy adatvesztés történne, ha akár csak egy adatot is kiszednék, illetve például a rangsor nem lenne értelmezhető az adott szezonban, ha egy egyént kiszednénk belőle.

Az adatokon egyetlen további változtatást végeztem még; a helyezés változót faktorrá alakítottam, mégpedig úgy, hogy felsőházas helyezésnek számít az első 8 helyezett – mivel ezek azok a csapatok, akik európai kupában indulhatnak a következő szezonban egy ilyen elért helyezés után. A középmezőnybe tartozik a 9-15. helyezett, és alsóházi csapatnak minősül a 16.-20. helyezést elérő, azaz kieső jelölt csapatok.

5.2 Panel elemzés

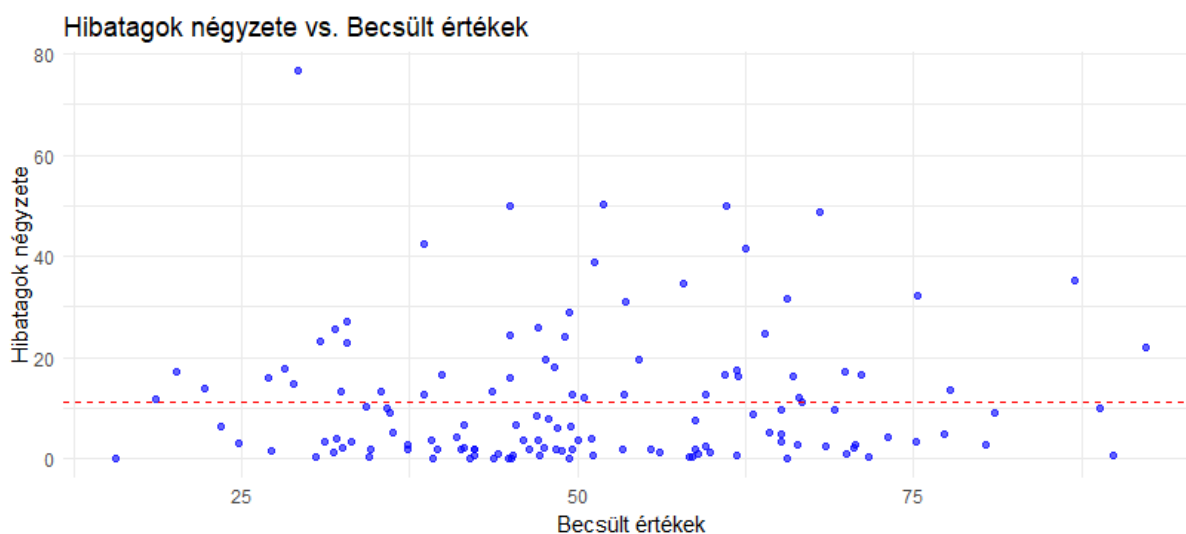
5.2.1 Alapmodell

Ahogy a módszertan fejezetben bemutattam, panel elemzést fogok végezni az adataimon még hozzá kiegyensúlyozatlanul, mert nem egyeznek meg a megfigyelt egyének a teljes adatbázisban a kieséses rendszernek köszönhetően. Az első modellek futtatásánál már egyből kijött, hogy a gólkülönbség túlságosan is függő változó, vagyis a modellben multikollinearitás van, ezért ezt a változót a továbbiakban elhagytam. A Chow-teszt eredménye

egy magas p-érték, 0,5819 lett, vagyis a nullhipotézist nem tudjuk elvetni, így a Pooled és fix modellek közül a Pooled-ot ajánlott választanunk. A Hausman-teszt alapján a p-érték (0,1862) megint magasabb, mint a szokásos 0,05-ös szignifikanciaszint, ami azt jelenti, hogy nincs elegendő bizonyíték arra, hogy a fix hatású és a véletlen hatású modellek eltérő eredményeket adnak. A Lagrange multiplikátor teszt eredménye szerint a p-érték (0,7065) szintén magas, ami azt jelenti, hogy a random modellel szembe is a Pooled az ajánlott modell. Ezen megállapítás alapján a Pooled OLS modellt kell használnom a továbbiakban. Először a modell heteroszkedaszticitását fogom megvizsgálni.

A heteroszkedaszticitás vizsgálata panelmodellek esetében is azért fontos, mert a heteroszkedasztikus hibatagok torzíthatják a standard hibákat, ami befolyásolja a t-statisztikákat és a szignifikancia szinteket. Ebben a konkrét esetben a heteroszkedaszticitás vizsgálata során a Breusch-Pagan tesztet alkalmaztam, amely nem mutatott szignifikáns heteroszkedaszticitást a modellben (p-érték = 0,7065). Ez azt jelenti, hogy az adatok hibatagjainak varianciája viszonylag állandónak tekinthető, és a modell becslései a klasszikus feltételezések szerint érvényesek maradhatnak. A hibatagok eloszlását az alábbi ábra mutatja:

Az ábra alapján nem jelenthető ki egyértelműen, hogy homoszkedasztikus lenne a hibatag, így a robusztusabb ellenőrzés érdekében HC1 és HC3 típusú robusztus standard hibákat is alkalmaztam, hogy biztosítsuk az eredmények stabilitását még akkor is, ha rejtett heteroszkedaszticitás lenne jelen. A robusztus standard hibák újraszámítása megerősítette a fő következtetéseket, és csökkentette a potenciális heteroszkedaszticitás hatását a t-statisztikákra és a szignifikanciaszintekre. Ezzel a módszerrel minimalizálhatók a standard hibák torzításaiból adódó problémák, így a modellünk eredményei megbízhatóbbá váltak.



4. ábra: Heteroszkedaszticitás ábrázolása - Forrás: Saját szerkesztés

A Pooled modell egyenletét az alábbiak szerint lehet felírni a regressziós együtthatók és változók alapján:

$$\text{Elért pontszám} = \beta_0 + \beta_1 \cdot \text{Középmezőny} + \beta_2 \cdot \text{Alsóház} + \beta_3 \cdot \text{Lőtt gólok} + \beta_4 \cdot \text{Kapott gólok} + \beta_5 \cdot xG + \beta_6 \cdot xGA + \dots + \beta_{28} \cdot \log xA + \epsilon$$

A becsült paramétereket behelyettesítve:

$$\begin{aligned} Pts = & 39,96 - 3,95 \cdot Rk_{\text{középmezőny}} - 10,37 \cdot Rk_{\text{alsóház}} \\ & + 0,36 \cdot GF - 0,46 \cdot GA + 3,04 \cdot xG - 2,92 \cdot xGA \dots + 3,27 \cdot \log xA + \epsilon \end{aligned}$$

Ebben az egyenletben a β_0 csak illeszkedést segítő paraméter, a β együtthatók a változók hatását becsülik a pontszámra (Pts), az ϵ pedig a hibatag. Az egyenlet mutatja, hogy a különböző független változók hogyan járulnak hozzá a pontszámhoz, amely a függő változónk.

A modell néhány együtthatójának értelmezését az alábbiakban ismertetem, a részletes eredményeket pedig a mellékletekben található táblázatok tartalmazzák (6. táblázat). A modell alapján a szezon során lőtt gólok pozitív hatással vannak a szezon végén szerzett pontok számára, míg a kapott gólok negatív hatást mutatnak, azaz a több lőtt gól összefügg a jobb szezon végi teljesítménnyel, a kapott gólok viszont csökkentik a pontszámot – ezek a megállapítások megegyeznek a szakirodalomban talált eredményekkel. A sportszakmai mutatók közül többnél is sikerült a kapcsolat irányát a várakozásoknak megfelelően kihoznom; az xG, a kapura lövések, passzpontosság és a labdabirtoklás is pozitív hatással van a csapat teljesítményére. Ellentétes irányú hatása volt az xGA-nak, pont úgy, mint ahogyan azt az előzetes várakozások mutatták.

A főbb gazdasági mutatókat nézve a tranzfer egyenleg és a tranzferköltségek pozitív, a magasabb bevételek pedig negatív irányú hatást mutatnak (utóbbi kettő logaritmálva volt). A tranzfer egyenlegre tett előzetes feltevés ezek alapján nem lett igazolva, sőt pont ellentétes eredmény jött ki, de a bevételekkel kapcsolatos bizonytalanságra kaptunk egy választ ezáltal. Végül a helyezést vizsgálva, a középmezőnyben elhelyezkedő csapatok eredményei alapján az eredmények azt mutatják, hogy ezek a csapatok kevesebb pontot szereznek a felsőházi csapatokhoz képest, ami alátámasztja a ligakon belüli teljesítménybeli különbségeket.

Megvizsgálva a változók szignifikanciáját, a szokásos szintekhez mérve csak az elért helyezés, a kapott és lőtt gólok, illetve a teljes piaci érték szignifikánsak. Az első változó egyértelműen adja magát, hiszen függvényszerű kapcsolat van a két változó között. A lőtt és kapott gólokat a szakirodalom bemutatásánál is említettem már, így azok magyarázó erejét sikerült reprodukálnom, ezzel ellentétben a piaci érték és a teljes nézőszám magyarázó ereje újdonság a vizsgált kutatásokhoz képest. Utóbbi szerint megállapíthatjuk, hogy a szurkolók

jelenléte ezek szerint befolyással van az eredményre, előbbi esetében pedig azt mondhatjuk, hogy a jobb csapatok, jobb játékoskerettel és emiatt jobb elért pontszámmal is fognak rendelkezni.

Modell szűkítése és standardizálása

A pontosabb becslés érdekében a modell szűkítése során elhagyok pár változót, melyek nem voltak szignifikánsak és a szakirodalom sem találta őket mind jelentős hatásúnak. Egészen pontosan csak a szignifikáns változók maradnak a modellben; ezek az elért helyezés, lőtt és kapott gólok, a teljes nézőszám, illetve a teljes piaci érték. A robosztus hibatagokat használó szűkített modell (ezzel a heteroszkedaszticitást kezeltem) értékeit az alábbi táblázat tartalmazza:

	Becslés	Standard hiba	t-érték	P-érték
<i>Konstans</i>	59,1693	4,5390	13,0357	0,0000
<i>Rkközépmezőny</i>	-4,0958	1,3775	-2,9734	0,0035
<i>Rkalsóház</i>	-10,1566	1,7165	-5,9171	0,0000
<i>GF</i>	0,4508	0,0497	9,0774	0,0000
<i>GA</i>	-0,4488	0,0332	-13,5053	0,0000
<i>Att_Total</i>	-0,0004	0,0002	-2,0660	0,0409
<i>Totalmarketvalue</i>	0,0078	0,0035	2,2454	0,0265

2. táblázat: Szűkített modell eredményei robosztus hibataggal - Forrás: Saját szerkesztés

A modell egyenlete:

$$\text{Elért pontszám} = \beta_0 + \beta_1 \cdot \text{Középmezőny} + \beta_2 \cdot \text{Alsóház} + \beta_3 \cdot \text{Lőtt gólok} + \beta_4 \cdot \text{Kapott gólok} + \beta_5 \cdot \text{Teljes nézőszám} + \beta_6 \cdot \text{Teljes piaci érték} + \epsilon$$

$$Pts = 59,17 - 4,09 \cdot Rkközépmezőny - 10,16 \cdot Rkalsóház + 0,45 \cdot GF - 0,45 \cdot GA - 0,0004 \cdot Att_total + 0,0078 \cdot Totalmarketvalue + \epsilon$$

A szűkített modell változóinak együtthatóit vizsgálva megállapítható, hogy a konstans változó itt is csupán az illeszkedést segítő paraméter. A modellem alapján a csapatok által szerzett gólok pozitív hatást mutatnak a szezon végi pontszámokra, míg a kapott gólok hatása negatív, azaz több lőtt gól általában a jobb teljesítménnyel, több kapott gól pedig a gyengébb eredményekkel hozható összefüggésbe. A modell további változói közül például a csapat piaci értéke szintén pozitív kapcsolatot mutat a teljesítménnyel, míg a teljes nézettség enyhén negatív irányú hatást jelez. Az eredmények alapján elmondható, hogy a szűkített modell együtthatóinak

irányai összhangban vannak az előző modellel, ami a következetes összefüggéseket támasztja alá.

Következő lépésként standardizáltam a modell folytonos változóit az összehasonlíthatóság érdekében. A mellékletek 7. táblázata alapján a legnagyobb (ráadásul pozitív) hatással az xGD változó van – egészen pontosan 55,3 az együtthatójának értéke. Így kijelenthető, hogy a várható gólkülönbségnek van a legnagyobb hatása a modellben a változóim közül. Ez újdonság a szakirodalomhoz képest, de ez betudható a statisztika újszerűségének. Ha összevetjük azzal, hogy a szakirodalomban említették, hogy a gólkülönbség egy alaptól fontos indikátor az elért pontszámhoz, akkor nem olyan meglepő az a következtetés, hogy a haladó statisztikai változata is jelentős hatású. Ezzel sikerült bizonyítani, hogy az újfajta statisztikai mutatók is számítanak a kutatási kérdés megválaszolásában. Igaz itt limitációként kezelni, hogy a változók közül mindössze a lőtt és kapott gólok voltak szignifikánsak.

5.2.2 Specifikus modellek

Az alapmodellemet a továbbiakban 2 újabb modellre fogom bontani; az egyikbe a hagyományos statisztikai mutatókat fogom használni, a másikban pedig az újfajta haladó statisztikákat. A modellek közös metszetei (például a gazdasági változók) egyfajta kontroll változókként fognak a modellekben szerepelni. A két modell fő különbsége a nem közös változókban rejlik; a hagyományos modell olyan teljesítménymutatókat tartalmaz, mint például a rangsor, a rúgott gólok, a kapott gólok, a passzpontosság, a lövések száma és a kaput eltaláló lövések aránya. Ezek olyan tényleges mérkőzésstatisztikák, amelyek közvetlenül a csapat teljesítményére vonatkoznak. Ezzel szemben az "expected" modell előre jelzett mutatókat alkalmaz, mint a várható gólok (xG), a várható kapott gólok (xGA), a várható gólkülönbség (xGD), a nem büntetéből származó várható gólok (npxG), a progresszív passzok, a progresszív cselek és a lövéshez vezető akciók (SCA). Mindezek a mutatók arra fókuszálnak, hogy milyen valószínűséggel ér el egy csapat jobb teljesítményt az előre jelzett statisztikák alapján. Mindkét modellben szerepelnek további közös kontrollváltozók, amelyek az általános csapattulajdonságokat és a pénzügyi helyzetet tükrözik; ezeket már fentebb részleteztem.

Először is megint el kell végeznem az alapvető modellválasztási lépéseket, vagyis megint le kell tesztelnem, hogy Pooled OLS, fixed effect vagy random effect modellre van szükségem a két modell esetében. Az eredmények itt is ugyanazt hozzák ki, mint az alapmodellnél; minden teszt p-értéke magas, vagyis a Pooled OLS modellt kell használni a továbbiakban is.

Teszt típusa	Modell	Chi-sq / F érték	Szabadságfok	p-érték
Hausman teszt	Trad	17,28	19	0,5709
F teszt (Chow)	Trad	0,93677	29, 84	0,5646
Lagrange (Breusch-Pagan)	Trad	1,2421	1	0,2651
Hausman teszt	Exp	23,051	21	0,3413
F teszt (Chow)	Exp	1,5355	29, 89	0,06535
Lagrange (Breusch-Pagan)	Exp	1,7047	1	0,1917

3. táblázat: Teszt eredmények modellválasztáshoz - Forrás: Saját szerkesztés

Az eredmények bemutatásánál már mindkettő modellnél robosztus standard hibákat használtam. A két új modell pontos eredményeit a mellékletek 8. táblázata tartalmazza. A hagyományos modell egyenlet az alábbi módon írható fel:

$$\text{Elért pontszám} = \beta_0 + \beta_1 \cdot \text{Középmezőny} + \beta_2 \cdot \text{Alsóház} + \beta_3 \cdot \text{Lőtt gólok} + \beta_4 \cdot \text{Kapott gólok} + \dots + \beta_{19} \cdot \log(\text{Átlagos játékosérték}) + \epsilon$$

A becsült paramétereket behelyettesítve:

$$\text{Pts} = 56,9717 - 4,1116 \cdot \text{Rk Középmezőny} - 10,5417 \cdot \text{Rk Alsóház} + 0,3916 \cdot \text{GF} - 0,4424 \cdot \text{GA} \dots - 5,76 \cdot \log \log \text{AVmarket_value} + \epsilon$$

A hagyományos modell becsült együtthatói alapján megállapítható, hogy a szezon során lőtt gólok pozitívan, míg a kapott gólok negatívan befolyásolják a csapatok szezon végi pontszámait, ami összhangban áll az alapmodell eredményeivel is. Ez a kapcsolat összhangban áll a szakirodalmi eredményekkel, amelyek szintén hangsúlyozzák a lőtt és kapott gólok jelentőségét a teljesítmény előrejelzésében. A labdabirtoklási arány is enyhe pozitív kapcsolatot mutat a szezon végi pontokkal, hasonlóan az alapmodellhez. Jelentősebb negatív kapcsolatot a helyezéseken túl a bevétel, a keretméret és az átlagos játékosértéknél fedezhetünk fel, mely valamennyire egyezik a saját előzetes feltevéseimmel, de nem lehet egyértelmű következtetést ezekből levonni a korábbi ellenkező eredményeim miatt sem.

Szignifikancia szintet vizsgálva ebben a modellben lett a legtöbb szignifikáns változóm a kutatás során. 5 százalékos szignifikanciaszinten szignifikánsnak számít a középmezőny és az alsó ház, a lőtt és kapott gólok, a teljes nézőszám, a teljes piaciérték és a logaritmált átlagos játékosérték. Ezek közül csak az utolsó jelent eltérést az alapmodellemhez képest – ezt a változót a logaritmálás miatt nem lehet szövegközben értelmezni, de az oka ugyanaz lehet, mint a teljes piaci értéknek, vagyis hogy a csapat teljesítményét erősen meghatározza a játékosok minősége.

Áttérve a haladó statisztikákat használó modellre; a mellékletek 9. táblázata alapján ennél a modellnél már egy kisebb eredmény is újdonságot hozhatna, ugyanis korábban a nem

standardizált modellekben a haladó statisztikai mutatók nem hoztak szignifikáns magyarázóerőt. A modell egyenlete:

$$\text{Elért pontszám} = \beta_0 + \beta_1 \cdot \text{Középmezőny} + \beta_2 \cdot \text{Alsóház} + \beta_3 \cdot xG + \beta_4 \cdot xGA + \beta_5 \cdot xGD + \dots + \beta_{22} \cdot \log xA + \epsilon$$

A becsült paramétereket behelyettesítve:

$$\text{Pts} = 62,11 - 7,07 \cdot \text{RkKözépmezőny} - 18,30 \cdot \text{Rkalsóház} - 2,67 \cdot xG + 2,59 \cdot xGA + 3,04 \cdot xGD + 2,93 \cdot xGA \dots + 1,96 \cdot \log xA + \epsilon$$

A haladó statisztikás modellnél már az egyenletre ránézve is több problémás megállapítás is tehető. Először az xG és xGA előjelei nem tűnnek ésszerűnek; a modell szerint a magasabb xG érték kevesebb ponttal jár, illetve fordítva az xGA esetében. Hasonló a helyzet a PPDA és SCA esetében; ha feltesszük, hogy progresszív játékkal lehet meccset nyerni, akkor ezeknél is értelmezhetetlen a negatív előjel. Ezek a megállapítások szembe mennek a sport alapjaival és a szakirodalmi eredményekkel is, így ezt a modellt nem is használok, szűkítem vagy értelmezem tovább, majd csak a standardizálással próbálkozom meg újra.

Specifikus modellek szűkítése

A modellszelekciót csak a hagyományos modellre fogom elvégezni. Abban a modellben a szignifikáns változók a középmezőny és az alsó ház, a lőtt és kapott gólok, a teljes nézőszám, a teljes piaciérték és a logaritmált átlagos játékosérték lettek így a szűkített tradicionális modellben ezeket fogom felhasználni. A modell eredményeit az alábbi táblázat tartalmazza:

	Becsülés	Standard hiba	t-érték	P-érték
Konstans	63,5973	4,5790	13,8889	0,0000
RkKözépmezőny	-4,0435	1,1118	-3,6369	0,0004
RkAlsó ház	-10,5255	1,4744	-7,1386	0,0000
GF	0,4591	0,0378	12,1372	0,0000
GA	-0,4510	0,0337	-13,3932	0,0000
Att_Total	-0,0004	0,0002	-2,1569	0,0329
Totalmarketvalue	0,0134	0,0040	3,3929	0,0009
log_AVmarket_value	-3,0550	1,6769	-1,8218	0,0709

4. táblázat: A szűkített hagyományos modell eredményei robosztus hibataggal - Forrás: Saját szerkesztés

A szűkített hagyományos modell egyenletét így lehet felírni:

$$\begin{aligned} \text{Elért pontszám} = & \beta_0 + \beta_1 \cdot \text{Középmezőny} + \beta_2 \cdot \text{Alsóház} + \beta_3 \cdot \text{Lőtt gólok} + \beta_4 \cdot \text{Kapott gólok} + \\ & \beta_5 \cdot \text{Teljes nézőszám} + \beta_6 \cdot \text{Teljes piaciérték} + \beta_7 \cdot \log(\text{Átlagos játékosérték}) + \epsilon \\ \text{Pts} = & 63,59 - 4,04 \cdot \text{RkKözépmezőny} - 10,53 \cdot \text{Rkalsóház} + 0,46 \cdot \text{GF} - 0,45 \cdot \text{GA} - \\ & 0,0004 \cdot \text{Att_Total} + 0,0134 \cdot \text{Totalmarketvalue} - 3,06 \cdot \log \text{AVmarket_value} + \epsilon \end{aligned}$$

A szűkített hagyományos modell együttthatói alapján a lőtt gólok változónak pozitív irányú hatása van a szezon végi pontszámokra, ami azt jelenti, hogy több szerzett gól összefügg a magasabb pontszámmal, kapott gólok hatása ezzel szemben negatív, ami összhangban van a korábbi modellekkel és a szakirodalmi megállapításokkal. A teljes piaci játékosérték továbbra is pozitív hatású, a teljes nézettség pedig enyhén negatív. A logaritmált átlagos játékosérték itt is negatív előjelet mutat, és csak a 10%-os szignifikanciaszinten lenne szignifikáns. Ez az eredmény az értelmezhetőség szempontjából bizonytalanabb, ami arra utalhat, hogy ennek a változónak a hatása a csapat teljesítményére nem egyértelmű, és további vizsgálatokat igényelhet. Összességében azonban a két modell együttthatói között csak minimális eltérés figyelhető meg, ami megerősíti a következtetések stabilitását.

Specifikus modellek standardizálása

Ahogy az alapmodell esetében, úgy a két új modellemre is elvégzem a folytonos változók standardizálását, hogy ezzel összehasonlíthatóbbak legyenek az eredményeim. Az eredmények alapján (mellékletek 10. táblázata) a hagyományos modell esetében a konstans után az alsóházbeli tagság (-10,54), a lőtt gólok (7,28) és a teljes piaci érték (6,98) a legmeghatározóbb együttthatók a modellben. A korábbiakhoz képest ennek az eredménynek az az újdonsága, hogy az alsó házhoz való tartozás jelentősen a legerősebb változó, ráadásul negatív irányba. Itt is limitációként kell kezelni, hogy a változók közül csak a már fentebb is említett pár változó szignifikáns, de a legerősebb magyarázóerejük közül mindegyik szignifikáns. A haladó modellem eredményeinek (mellékletek 11. táblázat) esetében a standardizálás után megállapítható megint ugyanaz, mint az alapmodell standardizálásánál, hogy kiemelkedően az x_{GD} a legmeghatározóbb együtttható a maga 66,42-es értékével. A konstans értéke után a következő két nagy értékkel rendelkező változó az x_G és az x_{GA} lettek, de itt megint előállt az az anomália, hogy fordítottan előnyös előjeleket kaptak, ahhoz képest, mint amit alapból fel lehetne tenni róluk.

6. Eredmények

A kutatás során végzett panel ökonometria elemzések lehetővé tették, hogy részletesen vizsgáljam a gazdasági és sportszakmai mutatók közvetlen és közvetett hatását a Premier League csapatainak teljesítményére. Az eredmények alapján meghatározhatóak azok a kulcsfontosságú tényezők, amelyek leginkább befolyásolják egy csapat szezonbeli sikerét. Az elemzés több modellből állt, amelyeket különböző változókra alkalmaztam: az első modell minden változót tartalmazott, amit a kutatásomhoz gyűjtöttem; a specifikus modelljeim közül

az egyik alapvetően a hagyományos statisztikai mutatókat használta, a másik a haladó statisztikákat, de voltak közös kontroll változók.

Az szakirodalmi eredményei alapján a Premier League klubjainak pénzügyi háttere, különösen a játékosbérek, statisztikailag szignifikáns kapcsolatot mutattak a csapatok teljesítményével, ám ezt nekem nem sikerült reprodukálnom az adataimon. Helyette sokkal meghatározóbb szerepet töltött be a csapatok teljes és átlagos játékos piaci értékei; minél magasabbak ezek az értékek (annál jobb minőségű játékosok vannak egy csapatban, vagyis) annál több pontot fog szerezni az adott csapat. A klubok pénzügyi teljesítményének elemzése továbbá rámutatott, hogy az átigazolásokra költött összegek kevésbé egyértelmű hatással bírnak az eredményváltozómra, ami az előzetes feltevéseimmel ellentétes.

A sportszakmai mutatók között különösen a támadáshoz kapcsolódó statisztikák játszottak jelentős szerepet a kutatásomban és az adatbázisomban is. Az elemzés során vizsgált mutatók közül különösen a lőtt és kapott gólok, illetve az xGD mutatott jelentős hatást – előbbiek szignifikanciában is konzisztensek voltak, utóbbi a standard modelleknél mutatott kiemelkedően magas együttható értéket. Továbbá több másik mutató kapcsolatának irányát is sikerült bizonyítani; ilyenek voltak a lövések, a labdabirtoklás, passzpontosság, az xG vagy az xGA. Utóbbi kettő esetében azonban meg kell jegyezni, hogy az az alapfeltevés, miszerint a nagyobb xG-vel rendelkező csapatok jellemzően több gólt szereztek és eredményesebb szezonokat zártak a haladó statisztikai modellemben pont fordítottnak jött ki, ezt a modellt nem is használtam fel, értelmeztem. A védekezési mutatók esetében az eredmények azt sugallták, hogy bár ezek a tényezők hatással vannak a csapat teljesítményére, a támadó mutatókhoz képest kisebb szerepet játszanak, az elemzés ezen eredményében valószínűleg az is szerepet játszik, hogy relatíve kevés védekezést leíró mutatót használtam fel.

Az eredmények alapján megállapítható, hogy a futballcsapatok teljesítményét jelentős mértékben befolyásolják mind a gazdasági, mind a sportszakmai mutatók. Az elemzés azt mutatta, hogy a klubok keretének értéke és a támadó mutatók (pl. szerzett gólok, xG) a legfontosabb tényezők a siker elérésében. A kutatás hozzájárul ahhoz a szakirodalomhoz, amely szerint a modern labdarúgásban a statisztikai elemzések és a gazdasági erőforrások szoros összefüggést mutatnak a szezonbeli eredményekkel, továbbá bizonyítja, hogy az újfajta statisztikai mutatóknak is hatásuk van a teljesítményre és releváns mutatóknak számítanak ezen a szakterületen.

7. Kutatási korlátok

A legfőbb kutatási korlátjaim az alapvető szignifikancia hiány a modellekben és az adatbázisom összeállítása; vagyis a korlátozottan elérhető gazdasági, pénzügyi mutatók voltak. Ahogy fentebb is említettem alapvetően a klubok transzferköltségei és a fizetési adatok voltak még a legegyszerűbben elérhető adatok. Ennek oka valószínűleg az elmúlt évek nagy érdeklődésének köszönhető az átigazolási piac körül, újságírók egy jelentős része arra húz fel manapság karriert, hogy átigazolási pletykákat és játékos szerződéseket szellőztet meg még a bejelentések előtt. A Premier League ilyen szempontból szerencsés választásnak bizonyult, mert a legtöbb ilyen adatot a klubok is meg szokták osztani, de más bajnokságokban az egyáltalán nem bevett szokás, hogy ezeket az adatokat nyilvánosságra hozzák, lévén, hogy üzleti titoknak minősülnek. Könnyen lehet, hogy az általam felhasznált adatok egy része is csak inkább becslés volt a hírek alapján, mint valódi megerősített információ, de minden egyes játékos adatainak utána menni nem volt célja a kutatásnak. Ugyanígy a pénzügyi adatokhoz tartozó korlát volt, hogy nem nyilvánosak a klubok részletes pénzügyeit összefoglaló adatbázisok sem. Egy nagyobb kutatás keretein belül a klubok saját oldalairól valószínűleg össze lehetett volna gyűjteni az éves beszámolókat, de az angol pénzügyi, számviteli beszámolási kötelezettségekről korlátozott információk álltak csak rendelkezésemre.

További korlátként megfogalmazható az is, hogy a vizsgált időszakom nem elég hosszú, így nem elég nagy a minta, illetve, hogy a vizsgált időszakba bele esett a Covid járvány okozta leállás és gazdasági válság is, mely egyértelműen befolyásolta a kutatást. A kutatás adatbázisában az első szezon, amely már bekerült a vizsgálatba alapvetően azzal magyarázható, hogy korábban még ne álltak rendelkezésre egyes mutatószámok, itt főleg a haladó statisztikai mutatókra kell gondolni.

A sportszakmai mutatóknál maradván korlátot jelentett még az is, hogy a választott sportágamban még a szakértők szerint sincsen egyértelműen olyan mutató, ami jól leírna egy csapat védekezési teljesítményét. Ez nagyrészt az eltérő stílusnak is köszönhető; például, ha valaki magasan meghúzott védelmi vonallal játszik, akkor az sok lesre állítással fog járni, de a kialakított helyzetek is magas xG-vel fognak rendelkezni, mert a nagyrészüik ziccerből jön, de valószínűleg kevesebb kapura lövés fog érkezni. Ezzel szemben, ha valaki „beparkolja a buszt” a kapuja elé, akkor nem lesz sok les, de mondjuk sok lövést fognak kapni, ezek minősége alacsony lesz és valószínűleg sokat is fognak blokkolni belőle – na és akkor hol az igazság, a két védekezési stílust a jelenlegi mutatókkal nehéz összehasonlítani.

8. Kitekintés, továbbfejlesztési lehetőségek

A korlátokat követően néhány továbbfejlesztési irányt fogalmaznék meg a jövőre való tekintettel. Maradva a védekezésnél; az utóbbi időben alakultak ki újfajta védekezést bemutató statisztikák, amelyeket kísérleti jelleggel meg lehet próbálni, hogy jól használhatóak-e. Ezek közé sorolható az expected threat vagy a kapusok teljesítményét vizsgáló védési hatékonyság, söprögetés esetleg a beadások lehúzásának hatékonysága, de fontos megjegyezni, hogy ezeknél is fennállhat az a korlát, amit a fentebbiekben említettem a védekezési stílussal kapcsolatban.

Nem csak a védekezést, de a játékosok egyéni teljesítményét is gyakran nehéz megítélni, még inkább nehéz összehasonlítani. Hogy lehetne összemérni egy alsóházi csapat kapusának a teljesítményét és bajnokesélyes csapat balszélsőjével például? A népszerű sporteredmény összegző oldalak használnak különböző értékelési rendszereket, de ezek közel sem megbízhatóak, és az egyes oldalak között még nagyobb az eltérés ilyen téren, mint a különböző xG modelleknél. Amennyiben lehetne valahogy standardizálni az egyéni teljesítményeket, úgy ezt a tényezőt is be lehetne vonni a modellbe.

Amennyiben az adatbázis vizsgált alanyait szeretnénk megváltoztatni, vagyis, hogy ne az angol férfi labdarúgó első osztály csapatait vizsgáljuk, úgy érdemes lehet más országok top bajnokságait is megvizsgálni, vagy kifejezetten alsóbb osztályokat elemezni (itt korlát lehet, az adatok elérhetősége), esetlegesen női bajnokságokra is elvégezni a fenti kutatást. Ezeket összevetve könnyen kijöhetnek különbségek, melyekkel következtetéseket meg lehet fogalmazni országokra, osztályokra vagy nemre vonatkozóan. Végül, de nem utolsó sorban; korlátként meg volt említve a vizsgált időszak is, mivel az magába foglalt egy Covid időszakot is – egy tisztább, béke időbeli adathalmazzal egyszerűbb és biztosabb megfigyeléseket lehetne összehozni.

9. Összefoglalás

A dolgozat célja az volt, hogy feltárja, mely sportszakmai és gazdasági mutatószámok jelzik előre a csapatok hosszú távú sikerét a Premier League-ben. Az elemzések során megállapítást nyert, hogy az egyes mutatók közötti összefüggések mélyebb megértése kulcsfontosságú a csapatok teljesítményének előrejelzésében és stratégiák kidolgozásában. A kérdés megválaszolásához először a szakirodalmi háttér volt fontos ismertetni. Ehhez először a sportgazdaságtani alapokat kellett lefektetni, hogyan működik egy liga, egyáltalán hogyan jött létre a Premier League; továbbá része volt a releváns elméletek és korábbi kutatások összefoglalása, amelyek meghatározzák a kutatási kérdésünk kontextusát. A szakirodalmi háttér

bemutatása során azokra a tanulmányokra összpontosítottam, amelyek az elemzett gazdasági és sportteljesítmény-mutatók hatását vizsgálták a csapatok hosszú távú sikerére vonatkozóan. Ezzel lett biztosítva, hogy a kutatás szilárd alapokon nyugodjon, és hozzájárulhasson a meglévő tudományos diskurzushoz.

Ezután a kutatási módszertan leírása következett, amely részletezte a vizsgálat során alkalmazott adatelemzési technikákat, az adatok forrását, valamint a statisztikai modellek kiválasztását. Az elemzéshez panelökonometria módszereket használtam, mivel az adatbázisom adottságai ezt diktálták. Az adatbázisom az elmúlt 7 szezon angol labdarúgó elsőosztály csapatainak adatait tartalmazta. Az adatok fő forrásai az Fbref, a Capology és Transfermarkt voltak. Első lépésként leíró statisztikai elemzést végeztem az adataimon, hogy amennyiben szükséges át tudjam alakítani a változókat; és néhányat az aszimmetriájuk miatt logaritmálnom is kellett.

A kutatás során több modellt is alkalmaztam, az első alapmodellem eredményei alapján csak az elért helyezés faktor változója, a kapott és lőtt gólok, a teljes nézőszám, illetve a teljes piaci érték rendelkeztek szignifikáns magyarázóerővel. A modell standardizálásával sikerült elérni, hogy bizonyítva legyen egy újfajta mutatószám magyarázóereje is; a standardizált modellben az xGD rendelkezett a legnagyobb együttthatóval. A továbbiakban 2 új modellt vezettem be; a hagyományos mutatókat és a haladó statisztikákat használt. Utóbbi nem bizonyult használhatónak, mivel több változónál is a várakozásokkal ellentétes hatást fedeztem fel. A standardizált verziója ennek a modellnek viszont újonnan ráerősített az xGD fontosságára és jelentőségére. A gazdasági mutatók közül a teljes piaci játékosérték bizonyult a legszignifikánsabb magyarázó változónak. A hagyományos modell esetén pedig sikerült nagyjából ugyanazokat a mutatókat szignifikánssá tenni, mint az alapmodellnél, és a változók értékei is hasonlóak lettek a standardizált verzióval. Az eredmények alapján kijelenthető, hogy néhány változónál sikerült reprodukálni a kapcsolatot, ami már a szakirodalmi tárgyalásnál is megjelent, továbbá sikerült bizonyítani pár újfajta mutató magyarázó erejét is. Ezek alapján megfogalmazható, hogy a gazdasági erőforrások és a statisztikai adatelemzés által nyújtott lehetőségek együttesen játszanak szerepet a csapatok sikerességében.

Fontos azonban rámutatni a kutatás korlátjaira: a vizsgált adatok és a használt módszertan alapján további mélyreható elemzések szükségesek lehetnek ahhoz, hogy pontosabb és általánosítható következtetéseket vonhassunk le. Ezen túlmenően, a sportág állandó változásai, például az új szabálymódosítások vagy a gazdasági helyzet alakulása szintén befolyásolhatják az eredményeket.

Összességében a dolgozat hozzájárul a labdarúgás területén folytatott kutatásokhoz azáltal, hogy hangsúlyozza az adatalapú döntéshozatal fontosságát és bemutatja, hogyan integrálhatók ezek a módszerek a gyakorlatba. Fő eredményeként kiemelhető az újfajta statisztikai mutatószámok közül néhány hasznosságának bizonyítása is. A jövőbeli kutatások számára érdemes lehet kiterjeszteni az elemzéseket más bajnokságokra is, valamint további változókat bevezetni a modellezésbe, hogy még szélesebb körű megértést érjünk el.

10. Forrásjegyzék

- Analyst, O. (2024.07.23). *Opta Football Stats Definitions*. Opta Analyst. <https://theanalyst.com/2024/07/opta-football-stats-definitions> Elérhető: 2024.11.11.
- András, K és Kozma, M (2015). Ligák nemzetköziesedése a hivatásos sportban. In: *Tanulmánykötet a 6. Báthory-Brassai nemzetközi konferencia előadásaiból*. Óbudai Egyetem Biztonságtudományi Doktori Iskola, Budapest, pp. 500-508. ISBN 978-615-5460-38-5
- András, K. (2004). A hivatásos labdarúgás piacai. 53. sz. *Műhelytanulmány*. Budapesti Közgazdaságtudományi és Államigazgatási Egyetem, Vállalatgazdaságtan Tanszék.
- Bács, B. A. (2020a). Az európai klubfutball meghatározó bajnokságainak összehasonlító elemzése. *International Journal of Engineering and Management Sciences*, 5(1), 22-35.
- Bács, B. A. (2020b). Futballóriások gazdasági elemzése= Economic Analysis of Football Giants. *Acta Carolus Robertus: Az Eszterházy Károly Egyetem Gyöngyösi Károly Róbert Campusának tudományos közleményei*, 10(2), 5-19.
- Beck, N., és Meyer, M. (2012). Modeling team performance: Theoretical and empirical annotations on the analysis of football data. *Empirical Economics*, 43(1), 335–356. <https://doi.org/10.1007/s00181-011-0463-2>
- Burton, C. (2022.06.01). *What is a free transfer in football? Bosman ruling explained*, Elérhető: <https://www.goal.com/en/news/free-transfer-football-bosman-ruling-explained/blt911adb4c29331153> letöltve: 2024. 06. 28.
- Catapult Sport Technology (2024). <https://www.catapult.com/> letöltve: 2024. 06. 28.
- Clemente, F. M. és Martins, F. M. L. (2017). Network structure of uefa champions league teams: association with classical notational variables and variance between different levels of success. *International Journal of Computer Science in Sport*, 16(1), 39-50. <https://doi.org/10.1515/ijcss-2017-0003>
- Collet, C. (2013). The possession game? A comparative analysis of ball retention and team success in European and international football, 2007–2010. *Journal of Sports Sciences*, 31(2), 123–136. <https://doi.org/10.1080/02640414.2012.727455>
- Deloitte Sports Business Group (2024.06.23.). Football Division Annual Review of Football Finance 2024, <https://www.deloitte.com/uk/en/services/financial-advisory/research/annual-review-of-football-finance-europe.html> Elérhető: 2024.11.11.

- Dendir, S. (2016). When do soccer players peak? A note. *Journal of Sports Analytics*, 2(2), 89–105. <https://doi.org/10.3233/JSA-160021>
- Ferri, L., Macchioni, R., Maffei, M., és Zampella, A. (2017). Financial versus sports performance: the missing link. *International Journal of Business and Management*, 12(3), 36. <https://doi.org/10.5539/ijbm.v12n3p36>
- Findikçi, M., és Tapşın, G. (2015). A panel data analysis of the relationship between league performance and the shares of the publicly-traded football clubs. *European Scientific Journal, ESJ*, 11(10). <https://eujournal.org/index.php/esj/article/view/6445>
- Franceschi, M., Brocard, J., Follert, F., és Gouguet, J. (2024). Determinants of football players' valuation: A systematic review. *Journal of Economic Surveys*, 38(3), 577–600. <https://doi.org/10.1111/joes.12552>
- Gásquez, R., és Royuela, V. (2016). The Determinants of International Football Success: A Panel Data Analysis of the Elo Rating. *Social Science Quarterly*, 97(2), 125–141. <https://doi.org/10.1111/ssqu.12262>
- Gómez, D. M., Holzmayer, F., Torgler, B., Schmidt, S. L., és Ortega, G. J. (2018). Transfer market activities and sportive performance in european first football leagues: a dynamic network approach. *Plos One*, 13(12), e0209362. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0209362>
- Göral, K. (2015). 2014 fifa dünya kupasının başarılı takımlarında pas başarı yüzdeleri ve topa sahip olma. *International Journal of Science Culture and Sport*, 3(9), 86-86.
- Griffiths, W. E., Hill, R. C., és Lim, G. C. (2012). Using Eviews: For principles of econometrics (4. ed). Wiley.
- Hobbs, J. (2021). *Measuring Transfer Success Through Goal Difference Added*. Analytics FC. <https://analyticsfc.co.uk/blog/2021/11/03/measuring-transfer-success-through-goal-difference-added/> Elérhető: 2024.11.11.
- Keleş, A. (2020). Examination of the relationship between the sources of income and the sports success of the clubs in the international football league. *Ankara Üniversitesi Beden Eğitimi Ve Spor Yüksekokulu SPORMETRE Beden Eğitimi Ve Spor Bilimleri Dergisi*, 37-45. <https://dergipark.org.tr/tr/pub/spormetre/issue/58354/656703>
- Lepschy, H., Woll, A., és Wäsche, H. (2021). Success Factors in the FIFA 2018 World Cup in Russia and FIFA 2014 World Cup in Brazil. *Frontiers in Psychology*, 12, 638690. <https://doi.org/10.3389/fpsyg.2021.638690>

- Liu, X. F., Liu, Y., Lu, X., Wang, Q., és Wang, T. (2016). The anatomy of the global football player transfer network: club functionalities versus network properties. *Plos One*, 11(6), e0156504. <https://doi.org/10.1371/journal.pone.0156504>
- Matsuoka, J., Packianathan, C., és Munehiko, H. (2003). Direct and interaction effects of team identification and satisfaction on intention to attend games. *Sport Marketing Quarter*, 12(4), 244–53.
- Modric T, Versic S, Drid P, Stojanovic M, Radzimiński Ł, Bossard C, Aftański T, Sekulic D. (2021). Analysis of Running Performance in the Offensive and Defensive Phases of the Game: Is It Associated with the Team Achievement in the UEFA Champions League? *Applied Sciences*. 2021; 11(18):8765. <https://doi.org/10.3390/app11188765>
- N. Harris, A. Miller (2011). Revealed: Official english football wage figures for the past 25 years. Elérhető: <https://www.sportingintelligence.com/2011/10/30/revealed-official-english-football-wage-figures-for-the-past-25-years-301002/> Elérhető: 2024.11.11.
- Naylor, A., és Harris, J. (2023). „A Cold War”: The rivalry between Brighton’s Tony Bloom and Matthew Benham at Brentford. The New York Times. <https://www.nytimes.com/athletic/3029279/2023/03/30/cold-war-brighton-tony-bloom-matthew-benham-brentford/> Elérhető: 2024.11.11.
- NSO (2017). PSG: eldőlt! Neymar ötéves szerződést irt alá – hivatalos <https://www.nemzetisport.hu/minden-mas-foci/2017/08/psg-eldolt-neymar-oteves-szerzodest-irt-ala-hivatalos> Elérhető: 2024.11.11.
- OPTA (2024). xG explained, <https://fbref.com/en/expected-goals-model-explained/>
- Phillips, B. (2011). Numberless Wonders. *Slate*. <https://slate.com/culture/2011/01/soccer-stats-prozone-opta-the-trouble-with-soccer-s-statistical-revolution.html> Elérhető: 2024.11.11.
- Pollard, R. (2002). Charles Reep (1904-2002): Pioneer of notational and performance analysis in football. *Journal of Sports Sciences*, 20(10), 853–855. <https://doi.org/10.1080/026404102320675684>
- Prayoga, H., Dharma, F., és Sukmasari, D. (2022). The effect of sports performance and financial performance on european soccer club stock prices. *Asian Journal of Economics and Business Management*, 1(2), 92-99. <https://doi.org/10.53402/ajebm.v1i2.83> Elérhető: 2024.06.11.
- Premier League (2024). <https://www.premierleague.com/> letöltve: 2024. 06. 28.

- Rocha-Lima, E. M., Tertuliano, I. W., és Fischer, C. N. (2021). The influence of ball possession, passes and shots on target in winning premier league football matches. *Research, Society and Development*, 10(8), e55110817824. <https://doi.org/10.33448/rsd-v10i8.17824> Elérhető: 2024.11.11.
- Rohde, M. and Breuer, C. (2016). Europe's elite football: financial growth, sporting success, transfer investment, and private majority investors. *International Journal of Financial Studies*, 4(2), 12. <https://doi.org/10.3390/ijfs4020012> Elérhető: 2024.11.11.
- Solhekol, K. (2024.04.30.). *Premier League spending cap rules explained: What might clubs be allowed to spend from the 2025/26 season?*, SKYSPORTS, Elérhető: <https://www.skysports.com/football/news/11095/13125832/premier-league-spending-cap-rules-explained-what-might-clubs-be-allowed-to-spend-from-the-2025-26-season> Elérhető: 2024.11.11.
- Statista (2023). Average annual salary of players in the Premier League in England in 2022/23, by club, Elérhető: <https://www.statista.com/statistics/675303/average-epl-salary-by-team/> Elérhető: 2024.11.11.
- Stefan Szymanski, 2022. "Is there evidence of softness in the budget constraint in football? Some evidence from English clubs," Chapters, in: Rasmus K. Storm & Klaus Nielsen & Zsolt Havran (ed.), *Professional Team Sports and the Soft Budget Constraint*, chapter 7, pages 155-171, Edward Elgar Publishing.
- Szabó, C., Szabó, P., és Soós, M. (2022). *Olvasd a játékot!* G-ADAM Kiadó.
- Thakkar, P., és Shah, M. (2021). An Assessment of Football Through the Lens of Data Science. *Annals of Data Science*, 8(4), 823–836. <https://doi.org/10.1007/s40745-021-00323-2>
- Tiippana, T. (2020). How Accurately Does the Expected Goals Model Reflect Goalscoring and Success in Football? <https://urn.fi/URN:NBN:fi:aalto-2020122758632>
- Tiwari, N. (2024.03.21.). *AI in Football? Liverpool FC Uses DeepMind's TacticalAI for High-Impact Corner Kicks.* Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2024/03/ai-in-football-liverpool-fc-uses-deepminds-tacticalai-for-high-impact-corner-kicks/> Elérhető: 2024.11.11.
- Torgler, B., és Schmidt, S. L. (2007). What shapes player performance in soccer? Empirical findings from a panel analysis. *Applied Economics*, 39(18), 2355–2369. <https://doi.org/10.1080/00036840600660739>

- Trail, G. T., Anderson, D. F., és Lee, D. (2017). A longitudinal study of team-fan role identity on self-reported attendance behavior and future intentions. *Journal of Amateur Sport*, 3(1). <https://doi.org/10.17161/jas.v0i0.5712>
- UEFA (2024.04.15.) The European Club Finance and Investment Landscape, https://editorial.uefa.com/resources/028c-1aa8afe670df-ba6bbce7e604-1000/20240415_circular_2024_16_en.pdf Elérhető: 2024.11.11.
- Užík, M., Warias, R., és Glova, J. (2022). Management of Transfer Prices in Professional Football as a Function of Fan Numbers. *Mathematics*, 10(16), 2982. <https://doi.org/10.3390/math10162982>
- Valenti, M., Scelles, N., és Morrow, S. (2020). Elite sport policies and international sporting success: A panel data analysis of European women's national football team performance. *European Sport Management Quarterly*, 20(3), 300–320. <https://doi.org/10.1080/16184742.2019.1606264>
- Wakefield, L.L., és Sloan, H.J. (1995). The effects of team loyalty and selected stadium factors on spectator attendance. *Journal of Sport Management*, 9(2), 153–172.
- Wanat, T. and Leksowski, Ł. (2022). Short- and long-term effects of transfers on football clubs' sport performance. *Journal of Education, Health and Sport*, 12(5), 252-260. <https://doi.org/10.12775/jehs.2022.12.05.019>
- Whitmore, J. (2023.08.08.). *Evolving Expected Goals (xG)*. Opta Analyst. <https://theanalyst.com/eu/2021/07/what-is-expected-goals-xg/>
- Wooldridge, J. M. (2013). Introductory econometrics: A modern approach (5th ed). South-Western, Cengage Learning.
- XFB ANALYTICS (2022) The History Of Football Analytics - Part 2. [Http://www.Xfbanalytics.Hu/Blog/Blog-Post/2](http://www.xfbanalytics.hu/blog/blog-post/2) Elérhető: 2024.11.11.
- Yaldo, L. and Shamir, L. (2017). Computational estimation of football player wages. *International Journal of Computer Science in Sport*, 16(1), 18-38. <https://doi.org/10.1515/ijcss-2017-0002>
- Zambom-Ferraresi, F., Rios, V., és Lera-López, F. (2018). Determinants of sport performance in European football: What can we learn from the data? *Decision Support Systems*, 114, 18–28. <https://doi.org/10.1016/j.dss.2018.08.006>
- Zivkovic J (2022). *_worldfootballR: Extract and Clean World Football (Soccer) Data_*. R package version 0.6.2, <https://CRAN.R-project.org/package=worldfootballR>

Adatok forrása

- Capology (2024.11.11.). Premier League salaries
<https://www.capology.com/uk/premier-league/salaries/> Elérhető: 2024.11.11.
- Fbref (2024.11.11.). 2017-2024 Premier League stats,
<https://fbref.com/en/comps/9/2023-2024/stats/2023-2024-Premier-League-Stats>
- Transfermarkt (2024.11.11.). Premier League Info, Income and expenditures,
<https://www.transfermarkt.com/premier-league/einnahmenausgaben/wettbewerb/GB1>
https://www.transfermarkt.com/premier-league/startseite/wettbewerb/GB1/plus/?saison_id=2023 Elérhető: 2024.11.11.

11. Mellékletek

<i>Rövidítés</i>	<i>Átlag</i>	<i>Szórás</i>	<i>Medián</i>	<i>Min</i>	<i>Max</i>	<i>Ferdeség</i>	<i>Csúcsosság</i>	<i>Standard hiba</i>
<i>Season_End_Year</i>	4,0	2,0	4,0	1,0	7,0	0,0	-1,3	0,2
<i>Squad</i>	14,8	8,9	14,5	1,0	30,0	0,2	-1,2	0,7
<i>Rk</i>	10,5	5,8	10,5	1,0	20,0	0,0	-1,2	0,5
<i>GF</i>	53,9	18,6	51,0	20,0	106,0	0,7	-0,1	1,6
<i>GA</i>	53,9	14,2	54,0	22,0	104,0	0,2	0,3	1,2
<i>GD</i>	0,0	29,6	-5,5	-69,0	79,0	0,6	0,1	2,5
<i>Pts</i>	52,6	18,6	49,0	16,0	100,0	0,5	-0,2	1,6
<i>xG</i>	51,9	13,7	49,4	28,8	92,0	0,8	0,2	1,2
<i>xGA</i>	51,9	11,3	52,6	23,8	78,0	-0,3	-0,2	1,0
<i>xGD</i>	0,0	22,6	-3,4	-43,2	64,0	0,7	0,0	1,9
<i>Attendance</i>	32156,8	20035,8	30817,5	178,0	74976,0	0,1	-0,8	1693,3
<i>Att_Total</i>	19167,0	3832,0	18291,0	11715	29961,0	0,7	-0,2	323,9
<i>Cmp_percent_Total</i>	77,6	5,2	77,4	61,2	88,7	-0,1	-0,4	0,4
<i>xA_Expected</i>	33,2	10,5	30,3	17,4	63,6	1,1	0,5	0,9
<i>Sh_Standard</i>	476,1	91,8	458,0	319,0	781,0	0,9	0,4	7,8
<i>SoT_percent_Standard</i>	33,2	2,9	33,3	26,9	40,9	0,0	-0,5	0,2
<i>npxG_Expected</i>	48,0	12,7	45,6	28,0	83,6	0,8	0,1	1,1
<i>Weekly_Wages£</i>	1616007,7	935376,9	1355990,5	386292,0	4593269,0	1,0	0,3	79053,8
<i>Age</i>	26,6	1,0	26,6	24,2	29,4	0,2	0,1	0,1
<i>Poss</i>	50,0	7,8	48,9	35,4	71,0	0,5	-0,6	0,7
<i>PrgC_Progression</i>	695,6	187,6	670,5	343,0	1325,0	0,8	0,8	15,9
<i>PrgP_Progression</i>	1462,4	382,3	1353,0	844,0	2570,0	0,8	0,0	32,3
<i>SCA</i>	844,4	175,3	804,5	567,0	1388,0	0,8	0,2	14,8
<i>PPDA</i>	11,9	2,5	11,9	6,4	19,0	0,3	-0,3	0,2
<i>Expenditure</i>	107,4	83,3	83,9	0,0	630,3	2,6	11,5	7,0
<i>Income</i>	48,2	51,2	31,7	0,0	278,3	1,6	2,8	4,3
<i>Balance</i>	-59,2	76,2	-51,0	-562,4	118,1	-2,1	12,0	6,4
<i>Squadsize</i>	39,9	5,3	39,0	29,0	56,0	0,6	0,2	0,4
<i>Foreigners</i>	24,3	5,2	24,0	11,0	40,0	0,1	-0,1	0,4
<i>AVmarket value</i>	12,2	7,7	10,1	3,0	40,6	1,1	0,8	0,6
<i>Total market value</i>	487,2	306,8	396,8	106,1	1460,0	0,9	-0,2	25,9

5. táblázat: Leíró statisztikai elemzés eredménye a változókra - Forrás: Saját szerkesztés

	<i>Becsítés</i>	<i>Standard hiba</i>	<i>t-érték</i>	<i>P-érték</i>
<i>Konstans</i>	39,9666	28,9448	1,3808	0,1703
<i>Rkközépmezőny</i>	-3,9492	1,7434	-2,2652	0,0256
<i>Rkalsóház</i>	-10,3712	2,1425	-4,8408	0,0000
<i>GF</i>	0,3642	0,0854	4,2640	0,0000
<i>GA</i>	-0,4558	0,0604	-7,5446	0,0000
<i>xG</i>	3,0378	10,5771	0,2872	0,7745
<i>xGA</i>	-2,9292	10,6381	-0,2753	0,7836
<i>xGD</i>	-2,9559	10,6351	-0,2779	0,7816
<i>Attendance</i>	0,0000	0,0000	0,2586	0,7964
<i>Att_Total</i>	-0,0013	0,0006	-2,0064	0,0474
<i>Cmp_percent_Total</i>	-0,1436	0,1860	-0,7722	0,4418
<i>Sh_Standard</i>	0,0394	0,0644	0,6119	0,5419
<i>SoT_percent_Standard</i>	0,2455	0,1870	1,3130	0,1921
<i>npxG_Expected</i>	-0,0649	0,3066	-0,2116	0,8328
<i>Age</i>	-0,0984	0,5044	-0,1950	0,8458
<i>Poss</i>	0,6902	0,4211	1,6392	0,1042
<i>PrgC_Progression</i>	-0,0013	0,0050	-0,2694	0,7882
<i>PrgP_Progression</i>	-0,0019	0,0043	-0,4400	0,6609
<i>SCA</i>	-0,0197	0,0343	-0,5725	0,5682
<i>PPDA</i>	0,4770	0,3087	1,5451	0,1254
<i>Balance</i>	0,0111	0,0175	0,6301	0,5300
<i>Squadsize</i>	-0,1726	0,1549	-1,1142	0,2678
<i>Foreigners</i>	0,0676	0,1170	0,5780	0,5645
<i>Totalmarketvalue</i>	0,0223	0,0060	3,7144	0,0003
<i>log_Weekly_Wages</i>	0,4636	2,2143	0,2094	0,8346
<i>log_Income</i>	-0,4858	0,6714	-0,7236	0,4710
<i>log_Expenditure</i>	0,4367	0,9350	0,4671	0,6414
<i>log_AVmarket_value</i>	-5,7597	3,2401	-1,7776	0,0784
<i>log_xA</i>	3,2665	5,4682	0,5974	0,5516

6. táblázat: Az alapmodell eredményei robusztus hibataggal - Forrás: Saját szerkesztés

	<i>Becslés</i>	<i>Standard hiba</i>	<i>t-érték</i>	<i>P-érték</i>
<i>Konstans</i>	52,6285714	0,3898641	134,992094	2,153E-129
<i>GF</i>	12,1752168	1,41094517	8,62912117	3,7746E-14
<i>GA</i>	-8,27388287	0,9279071	-8,91671465	8,128E-15
<i>xG</i>	-37,4308668	135,178611	-0,27689933	0,78235021
<i>xGA</i>	27,747278	111,621957	0,24858262	0,80412322
<i>xGD</i>	55,2829193	224,132188	0,24665319	0,8056126
<i>Attendance</i>	0,69559066	0,55113803	1,26209884	0,20944489
<i>Att_Total</i>	0,82923986	1,66763942	0,4972537	0,61995136
<i>Sh_Standard</i>	2,1685243	4,91759848	0,44097222	0,66005388
<i>npG_Expected</i>	2,2020981	3,45887609	0,63665134	0,52560632
<i>Age</i>	0,08635889	0,53610104	0,16108696	0,87230512
<i>PrgC_Progression</i>	0,37890752	1,14349491	0,33135917	0,74097107
<i>PrgP_Progression</i>	-1,90143144	1,58467298	-1,19988885	0,23262742
<i>SCA</i>	-1,89126035	5,1781606	-0,36523787	0,71559895
<i>PPDA</i>	-0,05750349	0,67442766	-0,08526264	0,93219957
<i>Balance</i>	-0,00627528	0,69947573	-0,00897141	0,99285736
<i>Squadsize</i>	0,41081087	0,70438217	0,58322156	0,56087633
<i>Foreigners</i>	-0,09753146	0,62219117	-0,15675481	0,87571059
<i>Totalmarketvalue</i>	2,25690355	1,8900106	1,19412216	0,23486633
<i>log_Weekly_Wages</i>	0,78454607	0,97419995	0,80532346	0,42228071
<i>log_Income</i>	-0,38124841	0,62247819	-0,6124687	0,54142648
<i>log_Expenditure</i>	-0,42841594	0,67913523	-0,63082567	0,52939544
<i>log_AVmarket_value</i>	-1,00287827	1,67564382	-0,59850325	0,55067108
<i>log_xA</i>	0,02950385	1,41551691	0,02084316	0,98340657

7. táblázat: Standardizált alapmodell eredményei - Forrás: Saját szerkesztés

	<i>Becslés</i>	<i>Standard hiba</i>	<i>t-érték</i>	<i>P-érték</i>
<i>Konstans</i>	56,9717	17,8225	3,1966	0,0018
<i>RkKözépmezőny</i>	-4,1116	1,3650	-3,0122	0,0032
<i>RkAlsó ház</i>	-10,5417	1,7351	-6,0756	0,0000
<i>GF</i>	0,3916	0,0577	6,7835	0,0000
<i>GA</i>	-0,4424	0,0381	-11,6131	0,0000
<i>Attendance</i>	0,0000	0,0000	0,5603	0,5764
<i>Att_Total</i>	-0,0011	0,0005	-2,3514	0,0204
<i>Cmp_percent_Total</i>	-0,0148	0,1628	-0,0912	0,9275
<i>Sh_Standard</i>	0,0069	0,0074	0,9319	0,3534
<i>SoT_percent_Standard</i>	0,2472	0,1387	1,7825	0,0774
<i>Age</i>	0,0997	0,3868	0,2578	0,7970
<i>Poss</i>	0,3108	0,2396	1,2971	0,1972
<i>Balance</i>	0,0098	0,0083	1,1798	0,2406
<i>Squadsize</i>	-0,1641	0,1277	-1,2849	0,2015
<i>Foreigners</i>	0,0664	0,0863	0,7691	0,4434
<i>Totalmarketvalue</i>	0,0228	0,0053	4,3280	0,0000
<i>log_Weekly_Wages</i>	0,2016	1,4332	0,1407	0,8884
<i>log_Income</i>	-0,4622	0,4592	-1,0066	0,3163
<i>log_Expenditure</i>	0,2653	0,5903	0,4495	0,6539
<i>log_AVmarket_value</i>	-5,7607	2,6306	-2,1899	0,0306

8. táblázat: Hagyományos modell eredményei robusztus hibataggal - Forrás: Saját szerkesztés

	Becslés	Standard hiba	t-érték	P-érték
<i>Konstans</i>	62,1092	26,6554	2,3301	0,0216
<i>RkKözépmezőny</i>	-7,0708	1,7082	-4,1394	0,0001
<i>RkAlsó ház</i>	-18,3028	1,9144	-9,5608	0,0000
<i>xG</i>	-2,6744	11,0815	-0,2413	0,8097
<i>xGA</i>	2,5965	11,1531	0,2328	0,8163
<i>xGD</i>	2,9337	11,1437	0,2633	0,7928
<i>Attendance</i>	0,0000	0,0000	0,5461	0,5861
<i>Att_Total</i>	-0,0008	0,0005	-1,6685	0,0981
<i>npG_Expected</i>	0,1749	0,2397	0,7297	0,4671
<i>Age</i>	0,5914	0,6528	0,9060	0,3669
<i>PrgC_Progression</i>	0,0017	0,0058	0,2971	0,7670
<i>PrgP_Progression</i>	0,0026	0,0042	0,6066	0,5454
<i>SCA</i>	-0,0109	0,0078	-1,4030	0,1634
<i>PPDA</i>	-0,1076	0,2944	-0,3655	0,7155
<i>Balance</i>	0,0170	0,0107	1,5847	0,1159
<i>Squadsize</i>	-0,2127	0,1494	-1,4240	0,1573
<i>Foreigners</i>	0,0600	0,1133	0,5292	0,5977
<i>Totalmarketvalue</i>	0,0295	0,0058	5,0558	0,0000
<i>log_Weekly_Wages</i>	-0,3875	2,0711	-0,1871	0,8519
<i>log_Income</i>	-0,9244	0,5589	-1,6539	0,1010
<i>log_Expenditure</i>	0,0892	0,6753	0,1322	0,8951
<i>log_AVmarket_value</i>	-3,6264	2,8099	-1,2906	0,1996
<i>log_xA</i>	1,9598	5,5208	0,3550	0,7233

9. táblázat: Haladó modell eredményei robusztus hibataggal - Forrás: Saját szerkesztés

	Becslés	Standard hiba	t-érték	P-érték
<i>Konstans</i>	33,8519	21,3946	1,5823	0,1164
<i>RkKözépmezőny</i>	-4,1116	1,1819	-3,4789	0,0007
<i>RkAlsó ház</i>	-10,5417	1,6198	-6,5081	0,0000
<i>GF</i>	7,2818	1,0004	7,2787	0,0000
<i>GA</i>	-6,2708	0,5419	-11,5710	0,0000
<i>Attendance</i>	0,2188	0,4462	0,4903	0,6249
<i>Att_Total</i>	-4,2969	2,2666	-1,8957	0,0606
<i>Cmp_percent_Total</i>	-0,0148	0,2101	-0,0707	0,9438
<i>Sh_Standard</i>	0,6312	0,8049	0,7842	0,4345
<i>SoT_percent_Standard</i>	0,2472	0,1787	1,3834	0,1693
<i>Age</i>	0,0993	0,4158	0,2387	0,8117
<i>Poss</i>	0,3108	0,2308	1,3464	0,1809
<i>Balance</i>	0,7475	0,5973	1,2515	0,2133
<i>Squadsize</i>	-0,8649	0,6583	-1,3139	0,1915
<i>Foreigners</i>	0,3465	0,4923	0,7039	0,4829
<i>Totalmarketvalue</i>	6,9828	1,9169	3,6427	0,0004
<i>log_Weekly_Wages</i>	0,1151	0,7440	0,1547	0,8773
<i>log_Income</i>	-0,6930	0,5009	-1,3834	0,1693
<i>log_Expenditure</i>	0,2277	0,5681	0,4008	0,6894
<i>log_AVmarket_value</i>	-3,5877	1,6459	-2,1799	0,0313

10. táblázat: Standardizált hagyományos modell eredményei - Forrás: Saját szerkesztés

	Becslés	Standard hiba	t-érték	P-érték
<i>Konstans</i>	59,2617	1,0488	56,5066	0,0000
<i>RkKözépmezőny</i>	-7,0708	1,5857	-4,4591	0,0000
<i>RkAlsó ház</i>	-18,3028	2,0145	-9,0855	0,0000
<i>xG</i>	-36,7394	151,2569	-0,2429	0,8085
<i>xGA</i>	29,2685	124,6238	0,2349	0,8148
<i>xGD</i>	66,4201	250,2954	0,2654	0,7912
<i>Attendance</i>	0,3567	0,6361	0,5608	0,5761
<i>Att_Total</i>	-2,9359	1,8363	-1,5988	0,1127
<i>npG_Expected</i>	2,2274	3,4350	0,6485	0,5180
<i>Age</i>	0,5887	0,5920	0,9944	0,3222
<i>PrgC_Progression</i>	0,3204	1,3443	0,2384	0,8120
<i>PrgP_Progression</i>	0,9825	1,8040	0,5446	0,5871
<i>SCA</i>	-1,9179	1,5484	-1,2386	0,2181
<i>PPDA</i>	-0,2676	0,7415	-0,3608	0,7189
<i>Balance</i>	1,2966	0,7861	1,6494	0,1019
<i>Squadsize</i>	-1,1211	0,8764	-1,2792	0,2035
<i>Foreigners</i>	0,3130	0,6977	0,4486	0,6546
<i>Totalmarketvalue</i>	9,0417	2,5176	3,5914	0,0005
<i>log_Weekly_Wages</i>	-0,2213	1,0813	-0,2047	0,8382
<i>log_Income</i>	-1,3860	0,6862	-2,0198	0,0458
<i>log_Expenditure</i>	0,0766	0,7708	0,0993	0,9211
<i>log_AVmarket_value</i>	-2,2585	2,0581	-1,0973	0,2749
<i>log_xA</i>	0,5702	1,5556	0,3665	0,7147

11. táblázat: Standardizált haladó modell eredményei - Forrás: Saját szerkesztés