

Eötvös Loránd Tudományegyetem

Társadalomtudományi Kar

Survey statisztika és adatanalitika MSc

Depressziós jelek azonosítása Facebook adatokkal

Egy kísérleti kutatás tapasztalatai

Témavezető:

Dr. Kmetty Zoltán

Egyetemi adjunktus

Készítette:

Lovák Zsanett

GST2DU

Survey statisztika és adatanalita szakos hallgató

Budapest, 2020

Köszönetnyilvánítás

Ezúton szeretnék köszönetet mondani konzulensemnek, Dr. Kmetty Zoltánnak, aki felkeltette érdeklődésem a téma iránt, a dolgozat elkészülése során végig segített, hasznos tanácsokkal látott el. Emellett szeretném megköszönni családomnak és barátaimnak a sok biztatást, motiválást, az egyetemi tanulmányaim alatt nyújtott rengeteg támogatást.

1.	<u>Bevezetés</u>	3
1.1.	<u>A téma motivációja</u>	3
1.2.	<u>A dolgozat felépítése</u>	4
2.	<u>Elméleti háttér</u>	6
2.1.	<u>A depresszió bemutatása</u>	6
2.2.	<u>A közösségi média hatásai</u>	7
2.3.	<u>Korábbi kutatások során használt módszerek és az eredmények ismertetése</u> .	7
3.	<u>A kutatás bemutatása</u>	16
3.1.	<u>Az adatgyűjtés folyamata</u>	16
3.2.	<u>Az adatbázisok bemutatása</u>	17
3.3.	<u>Etikai és jogi megfontolások</u>	18
3.4.	<u>Kutatási kérdések</u>	19
3.5.	<u>A függő és a kontroll változók létrehozása</u>	20
3.6.	<u>Elemzés</u>	27
3.6.1.1.1.	<u>Az éjszakai fokozottabb aktivitás vizsgálata</u>	27
3.6.1.1.2.	<u>Az időbeli aktivitás változékonyságának vizsgálata a bejegyzéseken keresztül</u>	30
3.6.1.1.3.	<u>Az időbeli aktivitás változékonyságának vizsgálata a reakciókon keresztül</u>	32
3.6.1.1.4.	<u>Az ismerősökkel való kapcsolat változása a tartalmakra adott reakciókon keresztül</u>	34
3.6.1.1.5.	<u>A gyakran említett témák vizsgálata</u>	37
4.	<u>Az eredmények értékelése</u>	42
5.	<u>Összefoglalás</u>	45
6.	<u>Irodalomjegyzék</u>	48
7.	<u>Melléklet</u>	51

1. Bevezetés

1.1. A téma motivációja

A depressziót napjainkban a leggyakoribb mentális betegségek között tartják számon. A betegség sokkal több embert érint, mint gondolnánk. Továbbá számos ismert tünete van, de pontos kiváltó okai nem ismertek, rengeteg tényező állhat a kialakulás hátterében. A mai felgyorsult világban minden egyes nap ki vagyunk téve valamilyen stresszfaktornak. A munkahelyi, az iskolai vagy akár az otthoni megfelelési kényszer, bizonyítási vágy nagy terhet jelent akár a gyerekek, akár a felnőttek számára. Ezt néhány ember könnyedén tudja kezelni, míg másokban ezeknek a stresszes időszakoknak, eseményeknek hatására kialakul egy torz önkép, állandó önvádaskodás, bizonytalanság, amely könnyedén hangulatzavarhoz, szorongáshoz vagy depresszióhoz tud vezetni. Emellett bizonyos gyógyszerek hosszú távú szedése szintén elősegíthetik a depresszió kialakulását.

A WHO 2020-as jelentése szerint [\(26\)](#) világszerte több, mint 264 millió ember szenved depresszióban. Az Eurostat felmérése szerint [\(10\)](#) 2014-ben az Európai Unió lakosságának közel 7%-a számolt be hosszabb ideje tartó depresszióról. Emellett Murray 1996-os tanulmányában [\(17\)](#) arról írt, hogy 2020-ra az iszkémiás szívbetegség (a szív rossz vérellátásából fakadó szívbetegség) után a major depresszió lesz a második leggyakoribb munkaképtelenséghez vezető betegség (1990-ben még a negyedik helyen állt). O'Connor és munkatársai (2009) [\(18\)](#) azt is megjósolták, hogy 2030-ra a három vezető betegség a HIV/AIDS, a depresszió és az iszkémiás szívbetegség lesz (utóbbi esetében vitatott, hogy az iszkémiás szívbetegség vagy a közúti balesetek állnak a harmadik helyen).

A fenti statisztikák arra mutatnak rá, hogy ez egy sokakat érintő, világszintű betegség, ennek ellenére az érintettek nem kapnak megfelelő segítséget. A mentális betegségek világszerte aluldiagnosztizáltak, ennek hátterében számos tényező áll. Az esetek túlnyomó részében a diagnózis a betegek saját bevallásán vagy a környezetük által jelentett mentális állapot vizsgálatán alapszik. Bár a páciensek kezelésére egyes országok biztosítanak alapellátási szolgáltatásokat, a depresszió felismerését és korai kezelését nehezíti, hogy a depresszióban szenvedők kevesebb, mint fele kér orvosi segítséget, azonban belül is többségük a háziorvoshoz fordul [\(25\)](#). A háziorvosok, sőt gyakran maguk a betegek sem ismerik fel a testi tünetekből a depresszió jelenlétét. A helyes diagnózist számos kivizsgálás, téves kórkép előzi meg. A szakemberek továbbképzése időigényes és költséges, így nincsen erre minden országnak megfelelő költségvetése. A sikeres kezelés mögött pszichiáterekből, klinikai szakpszichológusokból, szociális munkásokból és egyéb szakemberekből álló komplett csapat áll, de sokszor a beteg nehezen tud kapcsolatba lépni a megfelelő ellátási szervezettel.

A depresszió világszintű aluldiagnosztizáltsága és alulkezeltsége arra enged következtetni, hogy a meglévő szűrési módszerek nem megfelelőek a betegség korai felismeréséhez, így az eddigi eljárásokat kiegészítve új módszerekre van szükség. Szakdolgozatomban azt szeretném bemutatni, hogy a közösségi média adatai, kiegészítve a felhasználók által kitöltött kérdőívvel, megfelelőek lehetnek a depresszió jeleinek korai detektálására.

1.2. A dolgozat felépítése

Dolgozatom első nagy részében a kutatásom elméleti háttérét ismertetem. Ebben a részben a depresszióról, mint gyakori mentális betegségről, valamint tüneteiről írok röviden. Továbbá bemutatok néhány ismert betegséget, amik a depresszió kockázatát növelhetik. Ezt követően a közösségi média szerepére hívnám fel a figyelmet. A közösségi hálózat világszerte gyors ütemben növekszik, manapság szinte minden ember tagja legalább egy közösségi oldalnak. Ezekon a felhasználók tartalmakat tehetnek közzé, amik kifejezik az egyén aktuális hangulatát, viselkedési tulajdonságait. Az emberek ahogy a napi sikereikről, pozitív gondolataikról, élményeikről beszámolnak ismerőseiknek, úgy a kudarcokat, negatív, depresszív gondolatokat is megosztják másokkal. A tevékenységi adatok hosszú időre eltárolódnak, így a közösségi média adatokon keresztül megvizsgálhatjuk a felhasználók érzelmi, nyelvi jellemzőit, társadalmi kapcsolatait, ezek időbeli változásait, amik hasznosak lehetnek a depresszió tüneteinek azonosítására.

Az elmúlt években számos kutatás kihasználva a közösségi média ezen előnyeit, a mentális betegségek és az öngyilkos gondolatok korai felismeréséhez a felhasználók adatait tanulmányozták rövid szűrővizsgálatokkal kiegészítve. A dolgozatban ezek közül néhány kutatást mutatok be és a főbb eredményeket ismertetem. Ezen korábbi vizsgálatok a Twitter, a Facebook, a Weibo és a Reddit fórum felhasználóinak tevékenységét tanulmányozták. Egyes kutatások során a résztvevők előzetesen önbevalláson alapuló kérdőíveket, szűrővizsgálatokat is kitöltöttek, míg más kutatás az elektronikus orvosi nyilvántartási rendszerben (EMR) szereplő depresszióra utaló betegségek nemzetközi osztályozásának (ICD) kódjait használta kiegészítő eszközként.

A dolgozat második nagy részében a kutatásomról számolok be. A depresszió jeleinek azonosítására egy 2019 évi adatfelvételből származó adatbázist használtam. Az adatgyűjtést egy piackutató cég készítette, amely során az ország keleti részén élő magyarokat egy online kérdőív kitöltésére kérték fel. A kérdések egyfelől a szabadidős tevékenységeikre, zenei érdeklődésükre, a Facebook tevékenységükre, valamint a közösségi médiáról kialakult véleményükre, politikai nézeteikre irányultak. Másrészt a felmérésben a felhasználóknak a módosított PHQ-9 (Patient Health Questionnaire) szűrővizsgálat 9 kérdését is feltették, ezzel betekintést nyerve az egészségi állapotukra. Az adatgyűjtő- és feldolgozó tevékenységekről kapott részletes tájékoztatást, illetve beleegyező nyilatkozat aláírását követően a résztvevők letöltötték a Facebook profil archívumukat. A dolgozat ezen részében részletesen bemutatom a Facebook adatokból készült adatbázisokat, illetve a kutatási kérdéseket tekintem át. Ezt követően ismertetem a depressziós diagnózis változó, valamint a két depressziós skála változó létrehozásának módszerét.

A következőkben a korábbi kutatások eredményeiből származó depresszióra utaló viselkedési mintázatokat tesztelem az említett Facebook tevékenységi adatokon. Több jellemzőt is megvizsgáltam annak céljából, hogy megtudjam mennyire hozhatóak kapcsolatba a depresszióval. A különböző szakirodalmak eltérő véleményen voltak az időbeli posztolási mintázatok depresszióra gyakorolt hatásáról. Ezáltal a kísérletben résztvevőknél megvizsgáltam, hogy mutatkozik-e eltérés a depressziós és nem depressziós emberek napszakonkénti aktivitásában. Továbbá a legtöbb szakirodalom az aktivitás fokozott változását egyértelmű előrejelzőnek írta. Ezáltal megvizsgáltam két időszak

összehasonlításával a Facebook aktivitásuk változását a felhasználók által közzétett tartalmakon, illetve az ismerőseik által közzétett tevékenységekre adott reakciókon keresztül. Emellett megfigyeltem a tartalmakra adott reakciókon keresztül, hogy az ismerőseikkel való kapcsolatuk hogyan változott az idő múlásával. Végül a depressziós személyek által gyakran említett témákat vizsgáltam, mint a munka, vallás, spiritualitás, család, egészség. Ebben az esetben a Facebook által generált érdeklődési kategóriákon keresztül vizsgáltam a depresszió diagnózissal, illetve a depresszió skálákkal való korrelációjukat.

Szakdolgozatom utolsó részében a fenti vizsgálatok eredményeit értékeltem ki, illetve a végén egy rövid összefoglaló olvasható a kutatásról.

2. Elméleti háttér

2.1. A depresszió bemutatása

A major depressziót a leggyakoribb pszichiátriai betegségek között tartják számon. Az Amerikai Pszichiátria Társaság által kiadott Mentális zavarok diagnosztikai és statisztikai kézikönyve (DSM-IV.) szerint [\(1\)](#) akkor diagnosztizálják a páciens major depresszióval, ha az elmúlt két hétben a nap jelentős részében a következőkben felsorolt tünetek közül legalább öt tünetet tapasztalt. Ezen tünetek az alábbiak: érdeklődés és öröm elvesztése, levert hangulat, fáradtság, alvászavar, csökkent koncentrációs és gondolkozási képesség, bűntudat, önvád vagy értéktelenség érzése, jelentős fogyás vagy hízás, öngyilkossági gondolatok, öngyilkossági kísérletek. A diagnózis felállításához az első két tünet közül legalább az egyik szükséges. Amennyiben a szomorú hangulat érzése mellett a táplálkozási zavar, alvászavar, fáradékonyság, csökkent önértékelés, csökkent koncentráció- és döntésképeség, valamint reménytelenség érzése közül legalább még két tünetet tapasztal a beteg, akkor a DSM-IV kézikönyv alapján minor depressziót állapítanak meg nála.

A betegség pontos kiváltó okai nem ismertek, de korábbi kutatások kimutatták, hogy egyes betegségek, például a cukorbetegség, a stroke, vagy a szívbetegség növelhetik a depresszió kialakulásának kockázatát. Számos epidemiológiai vizsgálat [\(22\)](#) [\(27\)](#) azt is alátámasztotta, hogy a daganatos megbetegedések kapcsolatban állnak a depresszióval. A daganatos emberek körében az átlaghoz képest nagyobb arányban tapasztalhatók depresszióra utaló tünetek, illetve a depresszióban szenvedő egyéneknél gyakrabban alakul ki daganatos megbetegedés az átlagpopulációhoz képest.

A depressziónak számos testi és lelki tünete van. Testi tünet például a fáradtság, az alvászavar, valamint az étvágyzavar, mely esetében az érintettek általában étvágy csökkenést, ritkábban étvágy növekedést tapasztalnak. Lelki tünetek lehetnek a szomorúság, levertség, félelem, aggodalom, önvádaskodás, önutálat, reménytelenség, vagy bűntudat érzete, súlyosabb esetben halálvágy vagy öngyilkosság gondolata. A mentális zavarban szenvedők mindennapjaira is hatással van a betegség, hiszen csökkenhet aktivitásuk, koncentrációs zavarok léphetnek fel. Emellett gyakori jellemző, hogy elvesztik humorérzéküket, jó kedvüket, valamint bezárkózottság figyelhető meg náluk. Ezek a szimptómák nem csak saját maguk, de környezetük számára is felismerhetővé válnak. A depresszió korai felismerése és kezelése rendkívül fontos, hiszen a nem kezelt betegeknél a tünetek egyre súlyosabbá válhatnak, a fent említett öngyilkos gondolatok öngyilkossági kísérletekhez vezethetnek.

A WHO 2020-as beszámolója szerint [\(26\)](#) évente világszerte körülbelül 800 ezer ember lesz öngyilkos, ez a második leggyakoribb halálozási ok a 15-29 éves emberek körében. Fontos megemlíteni, hogy az öngyilkossági kísérletek és a befejezett öngyilkosságok hátterében 90%-ban elmondható, hogy valamilyen kezeletlen pszichiátriai betegség áll. Egyes közösségi oldalak már tettek előrelépések afelé, hogy olyan funkciókat hozzanak létre, amelyek hozzájárulhatnak ahhoz, hogy az érintett személyek gyorsan, megfelelő segítséget kapjanak, ezáltal megfékezve az öngyilkossági kísérlet bekövetkeztét. A Facebook automatikusan nem keresi az öngyilkosságra vagy mentális betegségekre utaló bejegyzéseket. Ahogy korábban már beszámoltam róla, általában az ismerősök, barátok

számára is felismerhetővé válik a probléma. A Facebook többek között erre a tényre alapozva, az öngyilkossági kísérletek megelőzésére egy olyan funkciót fejlesztett ki, ami alapján a felhasználók ismerősei bármikor jelenthetik, ha az adott illetőnél mentális vagy egyéb problémára utaló jeleket tapasztalnak. A bejegyzés melletti linkre kattintva, a "Támogatás kérés vagy bejegyzés jelentése" menüpontban kiválaszthatják az "Öngyilkosság vagy önsértés" problémáját. Ezután két opciót is felajánl számukra az oldal. Az elsőnél kérhetnek szakszerű öngyilkosság megelőzési tippet, hogy mit mondjanak ismerősüknek. A második lehetőség választása esetén a Facebook átvizsgálja a bejegyzést és az érintett személynek e-mailt küld, amiben segélyvonal hívására és szakember felkeresésére buzdítja.

2.2. A közösségi média hatásai

A közösségi hálózatok elterjedése világszerte folyamatosan növekszik. A Facebook jelenleg a világ legnagyobb közösségi hálózata, a hivatalos honlapon nyilvánosságra hozott adatok szerint 2020 második negyedében havonta több, mint 2,7 milliárd aktív felhasználóval rendelkeznek. Aktív felhasználóknak azok a személyek számítanak, akik az elmúlt 30 nap során bejelentkeztek a Facebook fiókjukba. Emellett a nyilvánosságra hozott adatokból az is kiderül, hogy közel 3,2 milliárd ember használta havonta valamelyik terméküket (Facebook, Instagram, Messenger, WhatsApp). Kijelenthető, hogy napjainkban azok a személyek, akik nincsenek jelen valamilyen közösségi hálón, a regisztrált felhasználókhöz képest kevesebb információhoz juthatnak vagy később juthat el hozzájuk az információ. A Facebook egyszerre biztosít szórakozást és informálódást is. A mai rohanó világban segít leküzdeni a társadalmi távolságokat, a felhasználók kapcsolatba tudnak maradni a világ bármely pontján barátaikkal, családjukkal, sőt a rég nem látott ismerősökkel is egyszerűen fel tudják venni a kapcsolatot. A Facebook felhasználók különféle média tartalmakat tehetnek közzé oldalukon vagy ismerőseik oldalán, amelyekben megoszthatják ismerőseikkel napi tevékenységüket, gondolataikat, érzéseiket, véleményüket bejegyzés, fénykép vagy videó formájában, továbbá megjelölhetik ezeken ismerőseiket, illetve reagálhatnak ismerőseik által közzétett tartalmakra. A közösségi oldalakon közzétett tartalmak kifejezik az egyének aktuális hangulatát és érzelmi állapotát. A felhasználó és ismerősei által közzétett nyilvános adatokból sok mindent megtudhatunk a felhasználó barátokkal és más közösségekkel folytatott társas interakcióiról. Továbbá ezen oldalak folyamatosan rögzítik a felhasználók nyelvi és viselkedési szokásait. A digitális háló segítségével passzív monitorozási folyamaton keresztül nagy mennyiségű longitudinális adatra tehetünk szert, miközben a felhasználók viselkedését természetes és nem laboratóriumi körülmények között vizsgáljuk. Számos kutatás foglalkozott a felhasználók által létrehozott közösségi média tartalmak alkalmazhatóságával a mentális betegségek kimutatására. A korábbi kutatások alátámasztották, hogy a szociális média adatok alapján információkhoz juthatunk a felhasználók társadalmi és pszichés állapotáról, így a meglévő szűrési eljárásokat kiegészítve segíthetnek a depresszió detektálásában.

2.3. Korábbi kutatások során használt módszerek és az eredmények ismertetése

A korábbi tanulmányok a közösségi média adatok gyűjtésére számos módszert mutattak be. Egyes kutatások során a résztvevőktől hozzáférést kértek közösségi média adataikhoz. Más kutatások a felhasználók nyilvános adatait használták elemzésük során. Nyilvános adatok származhatnak olyan mentális betegségekkel foglalkozó fórumokból is,

ahol a felhasználók megvitathatják egészségügyi állapotukat, betegségeiket, illetve tanácsokat kérhetnek másoktól, támogatást nyújthatnak másoknak. A kutatók ezeken az online fórumokon egyrészt a közzétett bejegyzéseknek a nyelvi elemeit vizsgálhatják, másrészt a nyilvános bejegyzésekben rákereshetnek olyan kulcsszavakra, ami arra utal, hogy az illetőt mentális betegséggel diagnosztizálták.

„Depresszióval diagnosztizáltak a múlt héten.”

Hasonló módszer, ha nyilvános bejegyzésekben a depresszióra vagy más mentális betegségre utaló jeleket keresnek.

„El szeretnék aludni és nem szeretnék soha felébredni.”

A nyilvános adatokon alapuló adatgyűjtéssel gyorsan és viszonylag olcsón nagy mintákra tehetünk szert. Fontos az ebből származó torzításokat is figyelembe venni, például valaki önmagát egy bejegyzésében tévesen diagnosztizálhatja.

A kutatók a depresszió felderítésére, előrejelzésére a közösségi média adatain kívül egyéb kiegészítő eszközöket is alkalmaztak, ilyenek az önbevalláson alapuló kérdőívek, szűrővizsgálatok (például CES-D, PHQ-9, stb.), illetve az elektromos orvosi nyilvántartási rendszerben (EMR) szereplő depresszióra utaló betegségek nemzetközi osztályozásának (ICD) kódjai. Az önbevalláson alapuló kérdőívek magas érvényességi és megbízhatósági szinttel rendelkeznek a depresszió kimutatására. Ugyan a klinikai interjúk rendelkeznek a legnagyobb érvényességgel és megbízhatósággal, de eddig ezeket egy tanulmány sem használta. A korábbi kutatásokban gyakran alkalmazták a depresszió kimutatásához az orvosok körében is elterjedt Beck Depressziós Kérdőívet, a PRIME–MD szűrővizsgálatot (Primary Care Evaluation of Mental Disorders - a mentális zavarok alapellátásbeli értékelése) vagy a CES-D szűrővizsgálatot (Center for Epidemiologic Studies Depression Scale). A Beck Depressziós kérdőív csoportosított állításokat tartalmaz. A betegnek egy csoporton belül a rá leginkább jellemző állítást kell bejelölnie.

„1. Nem vagyok szomorú.”

„2. Szomorú vagy nyomott vagyok.”

„3. Mindig szomorú vagyok, és nem tudok kikeveredni belőle.”

„4. Annyira szomorú és boldogtalan vagyok, hogy nem bírom tovább.”

A kérdőív nagy pontossággal előrejelzi a jól megválasztott határérték felett (19 pont) a major depressziós állapotot, így az orvos szükség esetén a beteget a pszichiátriai ellátás felé irányíthatja [\(20\)](#). A PRIME-MD 26 eldöntendő kérdésből álló kérdőív, amiből a páciens elmúlt hónapban tapasztalt tüneteiről kapunk információt.

„Az elmúlt hónapban gyakran zavarta-e az, hogy lehangolt, depressziós vagy reménytelen?”

A CES-D a depresszió tüneteinek mérésére alkalmas önértékelésen alapuló 20 kérdéses szűrővizsgálat. A válaszadónak arra kell válaszolnia, hogy az elmúlt egy hetes időintervallumban milyen gyakran tapasztalta a felsorolt érzésekre és viselkedésekre vonatkozó állításokat. A kérdőívben maximum 60 pont szerezhető. Általában az egyének

depressziós szintjét három osztályba sorolják: 0-15 pont az alacsony, 16-22 pont az enyhe vagy közepes, 23-60 pont a magas tartomány.

A szűrő kérdőívek a mentális betegségek kimutatásával kapcsolatban megbízható eredményeket mutatnak, reprezentatívak a populációra nézve. Azonban ezeknek az adminisztrációja időigényes, költséges, illetve válaszadói torzítások léphetnek fel, aminek számos oka lehet. Például a megkérdezettek nem szánnak elég időt a kérdőív kitöltésére, megfontolatlanul, nem őszintén válaszolnak a kérdésekre. Emellett a survey felmérések során egy adott pillanatban felkérlik a válaszadót, hogy hosszabb időszakra visszaemlékezve részletesen nyilatkozzon egészségi állapotáról, hangulatáról, ami szintén torzításokat eredményezhet. De Choudhury és munkatársai (2013) [\(4\)](#) kutatásuk során az átgondolatlan válaszok elkerülése érdekében két lépcsős ellenőrzést alkalmaztak. A CES-D szűrővizsgálat mellett a Beck Depressziós kérdőívet is használták. Azt feltételezték, hogy adott személy esetén a két tesztre kapott pontok hasonlóképp alakulnak, amennyiben a válaszadó megfontoltan tölti ki a teszteket. Továbbá kizárták azokat a válaszadókat, akiknek a kitöltési ideje két percnél rövidebb volt.

A korábbi kutatásokban a Twitter, a Facebook, az Instagram, a Weibo, illetve a Reddit fórum felhasználóinak tevékenységeit figyelték meg. A megfigyelt viselkedési jellemzők alapján építettek fel egy prediktív modellt. Ilyen jellemzők lehetnek például a felhasználók nyelvi jellemzői, amelyet a bejegyzésekben használt szavak gyakoriságán, a bejegyzések hosszán keresztül teszteltek.

De Choudhury és munkatársai 2013-as tanulmányukban [\(6\)](#) a Twitter felhasználók nyelvi és aktivitási jellemzőinek, valamint a társadalmi kapcsolati változásainak felhasználásával a súlyos depressziós rendellenességet, az MDD-t (Major Depressive Disorder) vizsgálták. A tömeges adatgyűjtésben az Amazon Mechanical Turk weboldalt használták, amely a világon 49 országban biztosít globális munkahelyet. A módszer segítségével gyorsan és költséghatékonyan gyűjthető információ különböző személyekről. Az MTurk segítségével a vállalkozók „tömegmunkásokat” alkalmazhatnak olyan online feladatok elvégzésére, amelyekre a számítógépek jelenleg nem képesek (például: fórumok vagy weboldalak moderálása, a felhasználók kérdéseinek megválaszolása, survey kérdőívek kitöltése, gépi tanulási modellek tanítása, stb.). A meghirdetett feladatokat HIT-eknek (Human Intelligence Tasks) nevezik. A kutatás során az Egyesült Államokban tartózkodó „tömegmunkásokat” kérték fel egy depressziós kérdőív kitöltésére. A felmérésben a résztvevők demográfiai adatairól, illetve depressziós kórtörténetéről gyűjtöttek információkat. A zajos válaszok elkerülése érdekében a korábban már említett két lépcsős szűrést alkalmazták, vagyis a CES-D szűrővizsgálat mellett a Beck Depressziós kérdőívet is használták a HIT-ben, illetve a két percnél rövidebb kitöltési idővel rendelkezőket kizárták. A feladat (HIT) elvégzésének száma egyre volt korlátozva, tehát egy személy csak egyszer tölthette ki a kérdőívet. A kitöltés végén a kutatók hozzáférést kértek a válaszadók Twitter adataihoz. A végső adatbázist azok a Twitter felhasználók alkották, akiket depresszióval diagnosztizáltak a kitöltést megelőző elmúlt egy évben, de a felmérés időpontja előtt legalább 3 hónappal korábban. Továbbá az elmúlt egy évben legalább két depressziós időszakon kellett, hogy áteszenek. A 476 felhasználó átlag életkora 25 év volt, a nemek aránya szinte egyenlő megoszlású volt (49% női, 50% férfi felhasználó). A felmérés alapján azokat tekintették MDD pozitívnak, akik legalább a CES-D szűrővizsgálat során 22 pontot

érték el.¹ Ezek alapján 171 személyt diagnosztizáltak súlyos depressziós rendellenességgel. Az MDD-vel diagnosztizált csoport esetén a betegség bejelentett megjelenésétől egy évre visszamenőleg, a másik csoport esetén a kitöltést megelőző egy évre visszamenőleg vizsgálták a Twitter tevékenységüket.

De Choudhury és munkatársai 2014-es kutatásukban [\(5\)](#) az előző tanulmányukhoz [\(6\)](#) hasonlóan a közösségi média adatokon keresztül a felhasználók társadalmi kapcsolatait, az érzelmi, nyelvi jellemzőit vizsgálták. Céljuk a szülés utáni depresszió (PPD=post-partum depression) korai felismerése volt. Az előzőtől eltérően, a résztvevőktől a Facebook adataikhoz kértek hozzáférést. Az anyák mentális állapotának felderítésére a Facebook tevékenységüket megfigyelték a szülés előtti 50 hetes időszakban, illetve a szülés utáni 10 hetes periódusban. Ezeket az adatokat a két időszak szerint elkülönítve, két adatbázisban tárolták. Olyan anyák vehettek részt a kutatásban, akik legfeljebb az elmúlt kilenc hónapban hozták világra gyermeküket és aktív Facebook profillal rendelkeztek. A résztvevők toborzásához a kisgyermekes édesanyákat célzó fórumokon bejegyzéseket, valamint levelezőlistákon, Facebook és Twitter oldalakon hirdetéseket tettek közzé. A résztvevőknek most is először egy kérdőívet kellett kitölteni, amelyben demográfiai adataikra, szüléssel kapcsolatos tapasztalataikra voltak kíváncsiak, beleértve a PPD tapasztalataikat. A diagnózis megismeréséhez a korábban említett CES-D szűrővizsgálat helyett a PHQ-9 kérdőívet (Patient Health Questionnaire) töltették ki velük. A PHQ-9 egy beteg egészségügyi kérdőív, ami az adott személy mentális állapotának gyors felmérésére szolgál, a DSM-IV által meghatározott depressziós tünetekre fókuszálva képes előrejelezni azt. A kérdőívben 9 kérdést tesznek fel, mindegyik az elmúlt két hét érzelmi, hangulati állapotára irányul (például: „Az elmúlt 2 hétben milyen gyakran okoztak Önnek gondot az alábbi panaszok? Kevés érdeklődés vagy örömezés tevékenységei során”). A válaszadó minden kérdésre egy pontszámot kap (0 - Egyszer sem, 1 - Néhány napig, 2 - A napok több, mint felében, 3 - Majdnem minden nap). A legtöbb vizsgálatban a depresszió súlyosságát úgy határozzák meg, hogy 0-4 pont esetén nem áll fenn depresszió, 5-9 pont esetén enyhe, 10-14 esetén közepes, 15-19 esetén közepesen súlyos és 20-27 pont esetén súlyos depresszióról beszélünk. Fontos megjegyezni, hogy a kérdőív még önmagában nem minősül orvosi diagnózisnak. Ismét próbálták a torzításokat elkerülni, ennek érdekében azokat a nőket kizárták a vizsgálatból, akik a PHQ-9 szűrővizsgálat alapján PPD-ben szenvednek, de saját elmondásuk szerint nem tapasztalták a betegséget. A végső adatkészlet (n=165) szerint 137 résztvevő a szűrővizsgálat és a saját bevallása alapján sem szenved szülés utáni depresszióban. 28 anya tapasztalta a PPD-t, ebből 13 személynek a PHQ-9 depressziós szűrőkérdőív is ezt az eredményt hozta. A demográfiai adatokat tekintve, a PPD-t tapasztalt édesanyák alacsonyabb jövedelemmel rendelkeznek, viszont az életkor alapján a két csoport között jelentősebb különbségek nem figyelhetőek meg.

Az előrejelző modell felépítéséhez mindkét esetben [\(5\)](#) [\(6\)](#) a felhasználók társadalmi kapcsolatait, érzelmi állapotát, nyelvhasználati jellemzőit vizsgálták. Emellett összehasonlították a depressziót tapasztalt és a nem tapasztalt személyek csoportját, annak megértésére, hogy milyen eltérő viselkedési mintázatokkal rendelkeznek. A szociális kapcsolatok feltérképezéséhez a felhasználók által közzétett bejegyzések, állapotfrissítések, megosztások, barátaiknak küldött tartalmak számát használták. Az eredmények azt

¹ Az MDD diagnosztizálásához számos tanulmány a CES-D szűrővizsgálat 16-os küszöbértéke helyett magasabb pontszámokat használ a hamis pozitív eredmények elkerülése érdekében.

mutatták, hogy súlyos depresszió, illetve szülés utáni depresszió keresztül ment személyek kevesebbet posztolnak, ezekre barátaik felől kevesebb reakció, komment érkezik, illetve az ismerőseik oldalára feltöltött bejegyzések száma leredukálódik. Összegezve, ahogyan az MDD-ben szenvedők esetén a Twitter, úgy a PPD-ben szenvedő anyáknál a Facebook aktivitás csökkenő tendenciát mutatott, inaktívabbak a közösségi oldalakon, barátaik kevésbé jeleznek nekik vissza. Ezek a viselkedési változások utalhatnak az illető bezárkózottságára, magányosságára, illetve a szociális kapcsolatok elvesztésére. Az édesanyák esetén fontos megjegyezni, hogy a fogyatkozó tartalmakból nem következtethetünk azonnal a depresszióra. Az édesanyák élete a szülés után egy új szakaszba lép, általában napjuk nagy részében a gyermekükkel foglalkoznak, így nem jut elég idejük a közösségi oldalakra.

Az érzelmi állapot mérésére a pozitív és a negatív érzelmi hatást vizsgálták. A pozitív és negatív érzelmi hatás esetében az LIWC (Linguistic Inquiry and Word Count)² szoftvert használták. Az érzelmi mérések során mindkét kutatásban a depresszióban szenvedők csoportjánál a negatív érzelmi hatás jobban megfigyelhető, fellelhetők a magány, szomorúság, lehangoltság, mint depressziós tünetek. A felhasználók bejegyzéseinek nyelvi jellemzőit is megfigyelték. Ez alapján mindkét kutatásban kimutatható volt, hogy a depressziós anamnézissel rendelkező személyek az egyes szám első személyű névmásokat gyakrabban használják. Ez kifejezi a társadalmi távolságtartást, valamint az önközpontúságot, miszerint önmagára nagyobb figyelmet fordít, mint környezetére. Továbbá elmondható, hogy a súlyos depresszióban, illetve a PPD-ben szenvedők többet káromkodnak a másik csoporthoz képest. A 2013-as kutatásukban [\(6\)](#) egy 1000 szavas depressziós lexikont telepítettek a Twitterre annak érdekében, hogy meghatározzák a felhasználók tweetjeiben a depresszió tüneteire utaló kifejezések gyakoriságát.³ Ráadásul a Wikipédia antidepresszánsok listáját felhasználva a bejegyzésekben a különböző gyógyszerek neveire is rákerestek. Megfigyelhető, hogy az MDD-vel diagnosztizált felhasználók a nem diagnosztizált csoporthoz képest bejegyzéseikben több depresszióhoz köthető kifejezést használtak. Ezek utalhatnak a lelki nyugtalanságra, a mentális instabilitásra. A bejegyzésekben a depressziós tünetek, a gyógyszerek gyakori említése arra mutat rá, hogy a közösségi médiában állapotukról, kezelésükről tartalmakat tesznek közzé. Emellett a munkával, családdal kapcsolatos aggályait is megvitatják, illetve a vallással és az ezotériával kapcsolatos témákban is gyakran tesznek közzé bejegyzéseket. Ezzel szemben a PPD-s édesanyák mentális és érzelmi állapotukra vonatkozó bejegyzéseket nem posztolnak, félve a barátok véleményétől, a társadalmi megbélyegzéstől. Az anyák az érzelmi és mentális állapotuk nyilvánosságra hozatalától elszigetelődnek, így az érzelmi állapot mérése nem mutatkozik jó PPD előrejelzőnek. A két depressziós betegségben szenvedő csoportban közös, hogy több kérdés tettek fel bejegyzéseikben, ami jelentheti azt, hogy a közösségi oldalakat információ gyűjtésére használják, vagy olyan személyek társaságát keresik, akikben ebben a helyzetben megbízhatnak, akik megértik őket. Megemlítendő, hogy a PPD-s anyák arról számoltak be, hogy általában a várt szociális, érzelmi támogatást nem kapták meg.

² LIWC szoftver: Az írott szöveg számítógépesített szöveganalízisének alapjául szolgál. Több éves tudományos kutató munka alapján határozták meg, hogy az egyes szavak a mindennapi nyelvezetünkben milyen gondolati, érzelmi, személyiségi és motivációs háttérrel rendelkeznek.

³ Ehhez a Yahoo!Answers Mentális Egészség kategóriájából válogattak hozzászólásokat (rövid kérdéseket és válaszokat). A leggyakrabban használt kifejezéseket a Wikipédiában a TFIDF numerikus statisztika segítségével kapták meg, amely megadja, hogy egy szó milyen sokszor fordul elő az egyes dokumentumokban. Ha eltérő hosszúságúak a dokumentumok, akkor gyakran végeznek súlyozást a helyes gyakoriság érdekében.

Eichstaedt és munkatársai (2018) kutatásuk során [\(8\)](#) szintén a Facebook közösségi oldalt használták a depresszió előrejelzésére. A korábbi tanulmányokban [\(5\)](#) [\(6\)](#) bemutatott önbevalláson alapuló kérdőívekkel, szűrővizsgálatokkal ellentétben, az elektromos orvosi nyilvántartásban (EMR) szereplő depresszióra utaló betegségek nemzetközi osztályozásának (ICD) kódjait használták a depresszió diagnózisának kiderítésére. A sürgősségi osztályra felvett betegektől (n=1175) hozzájárulást kértek a Facebook archívumuk letöltésére, illetve az EMR adataikhoz való hozzáférésre. A depressziós anamnézissel rendelkezők (n=114) esetében az első diagnózis dátumától kezdve vizsgálták a közösségi média adatokat, akár hat évre visszamenőleg. A korábban említett két kutatáshoz hasonlóan [\(5\)](#) [\(6\)](#), itt is megfigyelték a depressziós és nem depressziós személyek viselkedés mintázataiban fellelhető eltéréseket. Ennek érdekében a 114 depressziós anamnézissel rendelkező személy mindegyikéhez 5 kontroll beteget társítottak. Az időbeli összehasonlítás lehetősége miatt, a kontroll személyek esetében szintén a depresszióban szenvedő beteg első diagnózisának dátumától vették figyelembe a Facebook adatokat.

A két csoport aktivitásának tanulmányozására a korábban leírt módszereket használták, vagyis a bejegyzések hosszát, gyakoriságát, közzétételének időbeli megoszlását vizsgálták. Az eredmények azt mutatták, hogy a közzétett tartalmak hossza éves szinten átlagosan nagyobb a depresszióval diagnosztizált egyéneknél, viszont szignifikáns különbség nincsen a két csoport időbeli tevékenységi szokásaiban. Ezzel szemben De Choudhury és munkatársai (2013) tanulmányukban [\(6\)](#) arról számoltak be, hogy a Twitter felhasználók késő esti órákban, éjszaka aktívabbak a közösségi oldalon, ami tükrözheti azt, hogy alvászavarban szenvednek. További fontos eredménye a tanulmánynak, hogy a depresszió diagnosztizálását megelőző 6 hónap adataiból már nagy pontossággal lehet előrejelezni a depresszió jelenlétét. Megfigyelhető volt itt is, hogy akiket depresszióval azonosítottak több egyes szám első személyű névmást használtak. Az érzelmi nyelvi jellemzők vizsgálata során a magányra, szomorúságra, lehangolt érzelmi állapotra utaló szavak megfigyelhetőek voltak. Korábbi tanulmányokból ismert, hogy a depresszióval azonosított egyének többször szorulnak orvosi ellátásra a másik csoporthoz képest [\(24\)](#), így vizsgálták az orvosi ellátásra, testi panaszokra utaló szavak használatát is. Az eredmény alátámasztotta, hogy a felhasználók közül a depresszióban szenvedők gyakrabban írnak erre utaló szavakat bejegyzéseikben.

Korábbi szakirodalmak megerősítették [\(20\)](#), hogy a depresszió szorosan összefügg az öngyilkossággal. Az öngyilkosságról folytatott kutatások nehézsége, hogy gyakran az egyénekről az adatokat az öngyilkosság vagy öngyilkossági kísérlet bekövetkezése után gyűjtik, így az előrejelzésben torzítások léphetnek fel. Továbbá a társadalmi megbélyegzéstől félve az egyének nem jelentik, ha öngyilkos gondolataik támadnak, esetleg már kísérletet is tettek rá, így kevés adat áll rendelkezésre. De Choudhury és munkatársai (2016) kutatásukban [\(7\)](#) a Reddit fórum nyilvános bejegyzéseit vizsgálták az egyének mentális állapotának megismerésére. A kutatás célja annak előrejelzése, hogy a mentális egészség témákról folytatott társalgások résztvevői a későbbiekben mekkora eséllyel folytatnak öngyilkos gondolatokról beszélgetéseket. A Reddit egy olyan szociális média fórum, ahol az emberek új híreket, bejegyzéseket, képeket, linkeket töltenek fel és vitatnak meg. Bárki a regisztráció után tartalmakat tehet közzé, mások bejegyzéseit értékelheti, szavazhat rájuk pozitív vagy negatív minősítéssel. Továbbá a bejegyzések érdeklődések, témák szerint vannak rendezve vagy al-osztályokba vannak sorolva, amiket subredditnek neveznek. A regisztráció anonim módon is történhet (a személyes adatok megadása nem kötelező), ezért egy személy több

felhasználói profillal is rendelkezhet. Ebből kifolyólag a tanulmányban egy "felhasználó" alatt egy felhasználói profilt értenek.

A tanulmány során a mentális egészség subredditből (MH=Mental Health), illetve az öngyilkos-segélykérő subredditből (SW=Suicide Watch) gyűjtötték össze a bejegyzéseket. Az MH fórum több mentális egészségi problémával foglalkozik, míg az SW fórum az öngyilkosságot fontolgatók megsegítésére, lebeszélésére helyezi a hangsúlyt. Mivel a Reddit fórum szabályzatában szerepel, hogy az öngyilkossággal kapcsolatos tartalmak mindenképp az SW fórum felé irányuljanak, így az öngyilkos gondolatokkal kapcsolatos egyéb bejegyzések más subredditben feltételezhetően kis számban fordulnak elő. Két felhasználói csoport jött létre, akik a vizsgált időszak alatt mentális egészségügyi témákról beszélnek (MH fórumba közzétesznek bejegyzéseket), majd a vizsgált második időszak alatt nem tesznek közzé tartalmakat az öngyilkos-segélykérő subredditbe, vagyis feltételezhetően nem támad öngyilkos gondolatuk (n=440). A másik csoport, akik kezdetben a mentális egészség témákról folytatnak beszélgetést, viszont később az öngyilkos gondolatokról társalognak (SW fórumba közzétesznek bejegyzéseket) (n=440). Az eltérő jellemzőket keresték a két csoport között. Ehhez a korábbi tanulmányokban is használt módszereket alkalmazták. Megvizsgálták az igék, főnevek, határozószavak, névmások gyakoriságát a felhasználók nyelvi jellemzőinek megismeréséhez. Ezen túl, a bejegyzések számát, hosszát, a felhasználók aktivitási szokásait, az ezekben történő változásokat, valamint a szóhasználatukat is elemezték.

A kutatás eredményei alapján azok a személyek, akik a jövőben beszélgetnének az öngyilkosság gondolatáról, több igét, határozószót, de kevesebb főnevet használnak, mint azok a mentális egészségről társalgó felhasználók, akiknél a jövőben az öngyilkosság gondolata nem áll fenn. Utóbbi utalhat arra, hogy elvesztették érdeklődésüket bármely téma, dolog iránt. Ahogy korábban említettem, Eichstaedt és munkatársai (2018) kutatásukban [\(8\)](#) arra a következtetésre jutottak, hogy a közzétett tartalmak hossza átlagosan nagyobb a depresszióval diagnosztizált személyek esetében. A második vizsgált időszakban az SW fórumba közzétevő felhasználók bejegyzéseiről ez szintén elmondható. Továbbá ezek az írások tartalmilag nehezebben érthetőek, összefüggéstelenebbek voltak, ami jelzi a kognitív képesség, a nyelvi koherencia csökkenését. Az eddig említett tanulmányok [\(5\)](#) [\(6\)](#) [\(8\)](#) arról számoltak be, hogy a névmások használata kifejezi az egyén mentális jólétét, társadalmi kötelékeit. Általában az öngyilkos hajlammal veszélyeztetett csoport egyes szám első személyű névmásokat használ, ami az önközpontúságot, elzárkózást, érzelmi szorongást fejezi ki. Ez a megállapítás azoknál a Reddit felhasználóknál sincs másképp, akiknél a jövőbeli öngyilkosság kockázata fenn állhat. Ezen felhasználók aktivitásának csökkenése, ahogy az eddigi tanulmányokban [\(5\)](#) [\(6\)](#), úgy ebben is megfigyelhető volt. A kevesebb tartalom, azokra kapott kevesebb reakció most is tükrözi a társadalomtól való elzárkózást, magányt, illetve a társadalmi interakciók csökkenését. Továbbá a bejegyzéseik nyelvi jellemzőinek tanulmányozásánál fellelhető az egyes szám második és egyes szám harmadik személyű névmások, valamint a többes szám első személyű névmások ritka használata, ami szintén az előbb említett tüneteknek lehet a jele. A korábbi kutatásokhoz hasonlóan [\(6\)](#) [\(8\)](#) szóhasználatukban erősen megnyilvánul mentális állapotuk. Az SW fórumra író személyek gyakran használtak öngyilkos hajlamra utaló szavakat (például: meghalni, sírni, elhagyott), amiben a reménytelenség, lehangoltság, szorongás tünetei vehetők észre. Mindazonáltal olyan szavak is találhatóak a bejegyzésekben, amik reményt,

társadalmi kapcsolatot fejeznek ki (például társadalmi támogatásra, kezelésre, gyógyulásra utaló szavak).

Cheng és munkatársai (2017) tanulmányukban [\(4\)](#) azt vizsgálták, hogy a számítógépes nyelvelemzési módszerek felhasználhatóak-e a kínai Weibo⁴ felhasználók öngyilkossági kockázatának, érzelmi szorongásának, depressziós vagy stressz hajlamának felmérésére. Úgy vélték, hogy a felhasználók a közösségi oldalakon beszámolnak másoknak öngyilkossági gondolataikról, mentális állapotukról, így a klasszikus kérdőív mellett a nyilvános Weibo bejegyzések elemzése is alkalmas eszköz lehet a probléma korai felismeréséhez. A kutatásuk során a magas és alacsony öngyilkossági kockázatú emberek csoportját hasonlították össze.

De Choudhury és munkatársai (2016) kutatási [\(7\)](#) eredményével megegyezően, a magasabb öngyilkossági kockázattal rendelkező személyek gyakrabban használnak névmásokat, előljáró szavakat, többértelmű szavakat. Emellett azt találták, hogy ezek a felhasználók kevesebb igét használnak, ami nem egyezik meg De Choudhury és munkatársai (2016) eredményeivel [\(7\)](#). Tovább erősítették annak tényét, hogy az öngyilkosságra hajlamosabb embereknél magasabb az átlagos szószám, tehát hosszabb bejegyzéseket tesznek közzé. A nyelvi jellemzők vizsgálatánál olyan összefüggéseket találtak, amikről egyenlőre korábbi szakirodalmak nem számoltak be vagy eltérő következtetésre jutottak. Mégpedig, a depressziós vagy a stresszes felhasználók gyakrabban használnak többes szám második személyű névmásokat. Emellett a stresszes felhasználók gyakrabban, a magas öngyilkossági kockázattal rendelkezők ritkábban használnak többes szám harmadik személyű névmást. Az egyes szám első személyű névmások használata az öngyilkosság kockázatával és az érzelmi szorongással nem áll kapcsolatban. Ez a megállapítás az eddigi szakirodalmakban bemutatott eredményeknek ellentmond [\(5\)](#) [\(6\)](#) [\(7\)](#) [\(8\)](#). Fontos viszont megjegyezni, hogy ebben a kutatásban a kínai nyelvi jellemzőket vizsgálták, amelyben az egyes szám első személy kétértelmű jelentést hordoz. Ez a kétféle értelmezés nem csak a bejegyzést író személyre utalhat, hanem arra a közösségre is, amibe a felhasználó beletartozik, ezáltal udvariasabb és barátságosabb fényben tünteti fel a felhasználót. Ebből kifolyólag az egyes szám első személyű névmások használata a kínai nyelvben nem biztos, hogy az önközpontúságot jelzi és nem feltétlen utalhat az öngyilkosság vagy az érzelmi szorongás valószínűségére úgy, mint az angol nyelvben. A szóhasználat vizsgálata során újabb eltérő eredmények születtek. De Choudhury és munkatársai arról számoltak be [\(7\)](#), hogy azok a felhasználók gyakran használtak öngyilkos hajlamra utaló szavakat, akik elmozdulást mutattak a mentál higiéniai témákról az öngyilkos gondolatok felé, ilyen például a halálhoz kapcsolódó szavak. Az eredmények alapján a Weibo felhasználókról ez nem mondható el. Azoknál a személyeknél, akiknél fennáll az öngyilkos gondolatok-, kísérletek veszélye, kevésbé hajlamosak ilyen szavak használatára. Az eltérő eredmények oka lehet az eltérő módszer a kezelt és a kontroll csoport összevetésében. Cheng és munkatársai vizsgálatukban a magas és alacsony öngyilkossági kockázatú emberek csoportját hasonlították össze. De Choudhury és munkatársai vizsgálatuk során [\(7\)](#) a jövőre tekintve próbálták meghatározni a mentális egészség témájú subredditek résztvevőinek csoportjából, hogy kik azok, akik hajlamosak lehetnek az öngyilkosságra. Kutatásukban a mentális egészség subreddit fórumra íróknak csoportját hasonlították össze azokkal, akik később már az öngyilkos- segélykérő subreddit fórumra is írtak. Továbbá az eredmények azt mutatják, hogy

⁴ A Sina Weibo Kína legnépszerűbb blog weboldala, melynek több mint 300 millió regisztrált felhasználója van.

súlyos stressztől szenvedő Weibo felhasználóknál jellemzőbb a halálra utaló szavak használata, még akkor is ha sohasem gondolkodnak el az öngyilkosságon.

Moreno és munkatársai a 2011-es kutatásukban [\(16\)](#) a főiskolai hallgatók (n=200) nyilvános Facebook bejegyzéseit vizsgálták. A korábbi tanulmányok arról számoltak be [\(5\)](#), hogy a vizsgált célcsoportok a nyilvánosság előtt ritkán merik felvállalni problémáikat, félve a társadalmi megbélyegzéstől. Általában a fiatalok sem fordulnak szakemberhez, próbálják elkerülni, hogy állapotuk, problémáik napvilágot lássanak, így a betegségük későn vagy sohasem lesz diagnosztizálva. Moreno és munkatársai az elmúlt egy év Facebook bejegyzéseit vizsgálva próbálták a felhasználók állapotfrissítéseiben a depressziós tünetekre utaló jeleket fellelni. Továbbá a megjelenő depressziós szimptómák és a demográfiai vagy egyéb Facebook használati jellemzők közti összefüggéseket keresték. A bejegyzésekben a depresszióra utaló jeleket a DSM-IV (Mentális zavarok diagnosztikai és statisztikai kézikönyve) szerint meghatározott kritériumok alapján azonosították. A DSM-IV alapján a betegnek két hetes időintervallumban kell a korábban felsorolt tünetek közül legalább öt szimptómát tapasztalnia. Ezáltal a Facebook bejegyzésekben is a két hetes időszakban legalább öt olyan bejegyzésre volt szükség, ami a DSM-IV kritériumainak eleget tesz. Az adatgyűjtés során a felhasználók demográfia adatairól és a Facebook használati szokásairól, ismerősei számáról gyűjtöttek információkat, valamint az elmúlt egy év állapotfrissítéseit is lekérték. A depresszió jelenlétére utaló bejegyzéseket eltárolták teljes szöveggel, időbélyeggel ellátva, valamint az ezekhez tartozó megjegyzéseket szintén rögzítették.

Eredményeik alapján a főiskolai hallgatók gyakran tesznek közzé érzelmi állapotukra utaló tartalmakat, továbbá úgy találták, hogy akik ezekre kommentek formájában az ismerősöktől választ kapnak, gyakrabban merik megvitatni nyilvánosan az állapotukat. Ugyanakkor ez arra enged következtetni, hogy az ismerősök készségesen állnak rendelkezésre egy bajba jutott társuk megsegítésében. A bejegyzésekben leginkább fellelhető tünet a depressziós hangulat volt. A bűntudat, értéktelenség érzése gyakran megmutatkozott a tartalmakban, valamint a levertség érzése szintén kimutatható volt. Az érdeklődés elvesztésére, étvágyzavarra utaló jelek kevés bejegyzésben jelentek meg, valamint öngyilkos gondolatokra vonatkozó tartalmakat nem találtak. A korábban leírt kutatások eredményei az éjszakai aktivitásra vonatkozóan ellentmondásokba ütköztek. De Choudhury és munkatársai (2013) tanulmányukban [\(6\)](#) arról számoltak be, hogy a Twitter felhasználók éjszaka aktívabbak a közösségi oldalon, ami alvászavarra utalhat. Eichstaedt és munkatársai (2018) kutatásukban [\(8\)](#) ellenkező eredményeket kaptak. A vizsgált bejegyzésekben szintén gyakoriak voltak az alvásproblémára való utalások. Bár a megfigyelt főiskolai hallgatók alvási nehézségeire vonatkozó vizsgálatánál figyelni kell arra, hogy az ő esetükben az alváshiányra, kimerültségre utaló bejegyzések nem feltétlenül a depresszióra utalnak, hanem a késő esti tanulásnak is betudhatóak. Ennek elkerülésére az alvási problémákra vonatkozó bejegyzéseket külön kódolták. Az eredmények alapján a résztvevők 30%-a jelzett alvászavart úgy, hogy depresszióra utaló egyéb tünete nem volt. Az eddig bemutatott kutatási eredmények azt támasztották alá, hogy a depresszióban szenvedők közösségi média aktivitása csökkenő tendenciát mutat. Ebben a vizsgálatban úgy találták, hogy azok a személyek, akiknél a tünetek kimutathatóak, a vizsgálat időpontjához képest rövidebb időn belül többször frissítették állapotukat, a depressziós jeleket nem mutató felhasználókhoz képest.

3. A kutatás bemutatása

3.1. Az adatgyűjtés folyamata

Szaktervezésben használt adatbázisok egy 2019-ben készült adatfelvételtől származnak. A kutatás azonosítója: FK 128981 (OTKA)⁵, a kutatás vezetője: Dr. Kmetty Zoltán (<https://fbpilot.tatk.elte.hu>). A terepmunkát 2019 áprilistól szeptemberig egy piacutató vállalat végezte, akik a keleti országrészben élő magyarokat kérték fel a részvételre három ezer forint jutalom ellenében. A kutatásban 150 személy vett részt, akik legalább heti rendszerességgel használták a Facebook fiókjukat. A minta nem valószínűségi kvótaminta, az életkorra és a nemre vonatkozó kvótákkal.

A GDPR (General Data Protection Regulation = Általános adatvédelmi rendelet) szabályozás szerint a közösségi média oldalak kötelesek felhasználók számára biztosítani, hogy a regisztrációjuk pillanatától kezdve letölthessék az összes tevékenységükről tárolt adatmásokat. Ez magába foglalja a profiladatokat, a barátok, követők és követések listáját, az összes közzétett bejegyzést, történetet, fényképet, videót, hozzászólást, reakciót, kedvelt oldalak listáját és minden egyéb fiókhoz kapcsolódó tevékenységet. Ezenkívül a fizetési előzmények, a tartózkodási helyek, a keresési előzmények, a biztonsági adatok és még számos információ letölthető. A felhasználók kiválaszthatják a felsoroltak közül, hogy milyen időintervallumban, milyen adatokat szeretnének kiexportálni, illetve HTML vagy JSON formátumban szeretnék megkapni a letöltött információkat. Az elkészítési idő az adatok mennyiségétől függően néhány perctől akár fél napig is eltarthat, illetve az elkészült másolat csupán néhány napig tölthető le a felhasználók számára. A résztvevőket emiatt megkérték, hogy már a vizsgálat előtti napon kezdjék meg a Facebook archívumuk letöltésének folyamatát a kiküldött részletes leírás segítségével. A kutatásban az adatok a felhasználók regisztrációja óta lettek letöltve. A kutatók döntése alapján az előbb felsoroltak közül bizonyos információk nem lettek lekérve az adatexportálás során, mint például a ritkán használt funkciókból létrejövő adatok. Továbbá a felhasználóktól adatvédelmi okokból nem kérték el a túlzottan személyesnek minősülő adatokat, mint például az üzeneteket, a képeket, a videókat, a keresési előzményeket, a fizetési előzményeket vagy a biztonsági adatokat. A résztvevők a beleegyező nyilatkozat aláírását követően JSON formátumban személyesen a piacutató cég irodájában töltötték le a Facebook adataik másolatát. A kutatók az adatokat anonimizált formában kapták meg.

A kutatás résztvevői egy 15 perces online kérdőívet is kitöltöttek, ami számos demográfiai kérdést tartalmazott. Ebből kiderült, hogy a nemek megoszlását tekintve 113 nő és 37 férfi vett részt a kutatásban, átlag életkoruk 30 év volt. Továbbá a résztvevők közel 45%-a egyedülálló, valamint főként középosztálybelinek vallja magát. Emellett a kérdések a saját és más felhasználók Facebook tevékenységére, a közösségi médiáról és a politikáról kialakult véleményükre is irányultak. Egy nagyobb kérdésblokk a szabadidős tevékenységükkel, zenei érdeklődésükkel volt kapcsolatos. A survey vizsgálat mellett egy egészségügyi kérdőívet is kitöltöttek a felhasználók. A PHQ-9 egészségügyi kérdőív egy módosított változatát alkalmazták a depresszió súlyosságának megítéléséhez. A Facebook felhasználóknak szintén 9 kérdést tettek fel, viszont az elmúlt két hetes időintervallum

⁵ Országos Tudományos Kutatási Alapprogramok

helyett az elmúlt egy évre vonatkozóan. Ezenkívül a pontozás is megváltoztatásra került, hiszen a korábban említett 4 fokú skála helyett 7 fokút alkalmaztak.

Gondoljon vissza az elmúlt 1 évre. Ebben az időszakban milyen gyakran okoztak Önnek gondot az alábbi panaszok?	Skála						
	1	2	3	4	5	6	7
Kevés érdeklődés vagy örömezés tevékenységei során	1	2	3	4	5	6	7
Szomorúság, lehangoltság vagy reménytelenség	1	2	3	4	5	6	7
Nehezen tudott elaludni, éjszaka könnyen felébredt vagy túl sokat aludt	1	2	3	4	5	6	7
Fáradtság vagy kevés energia	1	2	3	4	5	6	7
Rossz étvágy vagy túlzott evés	1	2	3	4	5	6	7
Rossz érzések saját magával kapcsolatban, vagy olyan gondolatok, hogy Ön sikertelen, vagy csalódást okozott önmaga vagy családja számára	1	2	3	4	5	6	7
Koncentrációs nehézségek, például újságolvasás vagy tévénézés közben	1	2	3	4	5	6	7
Mozgása vagy beszéde annyira lelassult, hogy mások is észrevehették, vagy ellenkezőleg, olyan nyugtalan volt, hogy a szokásosnál sokkal többet mozgott	1	2	3	4	5	6	7
Olyan gondolatok, hogy jobb lenne meghalni, vagy hogy valamilyen módon kárt tehetne önmagában	1	2	3	4	5	6	7

1. Táblázat A módosított PHQ-9 kérdőív

3.2. Az adatbázisok bemutatása

Az adatok nyolc adatbázisban vannak tárolva, az első a kérdőívre adott válaszokat tartalmazza, a további hét a Facebook archívum adatait [\(13\)](#). A következőkben a hét adatbázist fogom bemutatni. A felhasználók ismerőseiről kapott adatok két adatbázisban lettek eltárolva, az egyik adatbázis a résztvevő azonosítóját, az ismerőseinek az anonimizált nevét, becsült nemét és az ismerőség kezdetét tartalmazza. A titkosítás előtt az ismerősök keresztnévét megvizsgálták és az alapján határozták meg a nemüket. A 150 résztvevő teljes ismerős listájának adatkészlete mintegy 116.000 személyt tartalmazott, amely közel 78.600 egyedi nevet jelent. Egy másik adatbázisban metaadatok szintjén állnak rendelkezésre az információk. Ez tartalmazza a válaszadó azonosítóját, ismerőseinek teljes számát, a függőben lévő és törölt kapcsolatainak számát, valamint az elutasított jelöléseknek a számát.

A Facebookon a közzétett tartalmakra a felhasználók többféle módon reagálhatnak, például kedvelhetik vagy viccesnek tarthatják az adott bejegyzést. Az adatokból nem ismerjük, hogy mi volt a bejegyzés pontos tartalma, de információt ad arról, hogy a felhasználó kinek a tartalmára, mikor és hogyan reagált. Emellett látható, hogy milyen típusú

közzétett tartalomra történt (például kép vagy komment) a reakció. A típusát tekintve 11 csoportba, plusz egy egyéb kategóriába sorolhatóak a reakciók, amelynek célpontja egy személy vagy egy oldal lehetett. Nagy mennyiségű adathalmazzal dolgozhatunk, hiszen több mint 1.800.000 reakciót tartalmaz az adatbázis. A mintában résztvevők 44,9%-ban bejegyzésre, 43,8%-ban képre reagáltak. A felhasználók leggyakrabban, 93,5%-ban a tetszik reakciót használták.

A következő adatbázisban a felhasználói tevékenységek 19 típusba, plusz egy egyéb csoportba lettek elkülönítve. Ilyen kategória lehet egy fotó vagy videó megosztása, bejegyzés közzététele valakinek az idővonalára, vagy épp egy cikkre történő komment írása. Az adatbázisból ismerjük a bejegyzés tartalmát, valamint ha érkezett rá megjegyzés vagy megosztott link, akkor annak pontos számát is tudjuk. Továbbá arról is van információnk, hogy a felhasználó melyik ismerősének idővonalára vagy milyen csoportba tett közzé bejegyzést. Természetesen ebben az esetben is csak az ismerősök anonimizált nevét ismerjük. A közel 350.000 időbélyeggel ellátott felhasználói tevékenység 26%-a valakinek az idővonalára közzétett bejegyzésből származik, 23%-a pedig a Facebook játékokból.

A felhasználók által követett oldalak szintén kategorizálva vannak, 19 típust különítettek el. Ebben az esetben az oldal típusa mellett a nevét is ismerjük, illetve az oldal követésének kezdeti időpontját. Az adatbázisban közel 52.400 követett oldal neve található.

Az utolsó két adatbázis a hirdetésekkel kapcsolatos adatokat tartalmazza. Az egyik adatbázis olyan válaszadói azonosítóval, időbélyeggel ellátott hirdetések listája, amik a felhasználó számára megjelentek a Facebook oldalán. Ez több mint 4200 egyedi hirdetést tartalmaz. A Facebook minden felhasználót érdeklődési kategóriákba sorol és ez alapján mindenkinek egyénre szabott reklámokat ad. Ennek folyamatáról pontos információnk nincsen, de valószínűleg a bejegyzések, kedvelt oldalak, keresési előzmények, ismerősök kedvelései alapján csoportosít. A másik adatbázisban a Facebook által generált hirdetési kategóriák listája szerepel, amely több, mint 18.600 egyedi érdeklődési kategóriába csoportosult.

3.3. Etikai és jogi megfontolások

A közösségi média alapú adatgyűjtés számos etikai és jogi kérdést felvet. Ezen adatok olyan információval bírnak az egyénekről, ami számukra érzékeny lehet, hiszen ahogy a korábbi tanulmány is alátámasztotta [\(5\)](#), a résztvevők félnek a mentális betegségek okozta társadalmi megbélyegzéstől, ezáltal az adatok jogszerű kezelése elengedhetetlen. Magyarországon a GDPR törvénye egyértelműen meghatározza az adatvédelmi irányelveket. Ez alapján a felhasználó hozzájárulása szükséges a személyes adatai kezeléséhez, csak a szükséges adatok kezelésére kerülhet sor, valamint célhoz kötöttnek kell lennie az adatkezelésnek. A 2019-es kutatás során a résztvevők részletes tájékoztatást kaptak az adatgyűjtő és adatfeldolgozó tevékenységekről. Adataik az írásos beleegyezésüket követően

lettek kiexportálva. Az adatfelvételt követően a kutatók részére az adatok anonimizálva kerültek átadásra.

3.4. Kutatási kérdések

Szakedolgozatomban a depresszió jeleinek azonosításához a Facebook archívumból származó tevékenységi adatokat, valamint a válaszadók által kitöltött online kérdőívet használtam. A depressziónak számos ismert tünete van, úgy mint az alvászavar, étvágyzavar, csökkent koncentrációs képesség, lehangoltság, reménytelenség, önutálat, félelem érzete, valamint az érdeklődés elvesztése vagy az öngyilkosság gondolata. A korábban bemutatott szakirodalmakban a kutatók a felhasználók nyelvi és aktivitási jellemzőinek, valamint a társadalmi kapcsolati változásainak felhasználásával azonosították ezeket a szimptomákat. Ezen eredményekben bemutatott viselkedési mintázatokat szeretném megvizsgálni az adatfelvételben résztvevő 150 ember adatainak esetében. Szakedolgozatomban a depressziós jelek azonosításához a Facebook közösségi oldal használatából kikövetkeztethető viselkedési mintákra fókuszáltam, a nyelvi jellemzők vizsgálatától eltekintettem. A nagy mennyiségű adatok lehetőséget biztosítanak ezen mutatók vizsgálatára. A kutatásom során az Excel program, illetve a statisztikai számításokhoz és az ábrázolásokhoz az R programnyelv volt a segítségemre.

Amint korábban arról beszámoltam Eichstaedt és munkatársai (2018) tanulmányukban [\(8\)](#) arra a megállapításra jutottak, hogy nincsen a két csoport időbeli tevékenységi szokásaiban szignifikáns különbség, viszont De Choudhury és munkatársai (2013) [\(6\)](#) szerint a depresszióval azonosított személyeknél az éjszakai fokozottabb közösségi média jelenlét megfigyelhető, ami arra utalhat, hogy alvászavarral küzdenek. Ezáltal az első kutatási kérdésem az volt (K1), hogy van-e eltérés a depressziós és nem depressziós emberek napszakonkénti aktivitásában. Ennek teszteléséhez megvizsgáltam a 150 felhasználó 2019 év első hat hónapra vonatkozó, Facebook tevékenységből származó adatait.

A szakirodalmak arról számoltak be [\(5\)](#) [\(6\)](#) [\(7\)](#), hogy a depresszióban szenvedő személyek esetében gyakori a közösségi oldalakon az aktivitás nagy mértékű változása. Előfordulhat, hogy egyes időszakokban rengeteg tartalmat tesznek közzé, míg a következő periódusban eltűnnek a közösségi oldalakról, ami arra utalhat, hogy bezárkóznak, megszakítják a szociális kapcsolatokat. Ebből adódóan a második kutatási kérdésem az volt (K2), hogy mutatkozik-e a depresszióval nem diagnosztizált személyekhez képest, a depresszióban szenvedő személyeknél nagyobb aktivitási változékonyság. Ennek megfigyeléséhez először a Facebook tevékenységi adatok közül a bejegyzések közzétételére vonatkozó adatokat használtam. A 2018 és a 2019 év első fél évében a bejegyzéssel töltött napok számát vetettem össze. Mivel a felhasználók gyakran reagálnak a közzétett bejegyzésekre, képekre és egyéb tartalmakra, így a bejegyzések vizsgálata után a résztvevők reakcióval töltött napjainak számát is hasonlóképp teszteltem.

A bejegyzéseken, reakciókon keresztül megfigyelt aktivitás változékonyságán túl, az ismerősökkel való kapcsolattartásban is megnyilvánul a felhasználók viselkedésének változása. A depresszióban szenvedő személyeknél a szociális visszahúzóds megfigyelhető, ami a közösségi oldalakon az ismerősökkel való interakciókon keresztül szintén észrevehető. A szakirodalmak eredményeire alapozva [\(5\)](#) [\(6\)](#) [\(7\)](#) a harmadik kutatási kérdésem az volt (K3), hogy változnak-e a depresszióban szenvedő személyek társadalmi kapcsolatai, valóban csökkenő tendenciát mutatnak-e. A felhasználók szociális kapcsolatainak változását a közzétett tartalmakra adott reakcióik alapján vizsgáltam. A korábbiakhoz hasonlóan a 2018 és 2019 első hat hónap adatát vetettem össze.

A mentális betegségben szenvedő személyekre jellemző, hogy szociálisan visszahúzódnak, de mégis a közösségi oldalakon gyakran megbeszélik problémáikat, valamint mindennapi küzdelmeiket a hasonló problémákkal küzdő személyekkel próbálják megvitatni [\(7\)](#). A korábbi kutatási eredmények szerint [\(6\)](#) [\(8\)](#) [\(16\)](#) a depressziós személyeknél a közösségi platformjaikon megfigyelhetők egyes témák gyakori említése. A családdal, munkával kapcsolatos problémáikat szeretik a közösségi oldalakon megvitatni, illetve többen írnak a spiritualitásról, valamint a vallásról. Több negatív hangvétélű, depresszióhoz köthető kifejezést használnak a depresszióban nem szenvedő személyekhez képest. Ezek a depressziós tünetekre, kezelésükre, sőt akár még gyógyszereikre is utalhatnak. A negyedik kutatási kérdésem az volt (K4), hogy gyakrabban említik-e a depressziós személyek ezeket a témákat, mint a depresszióval nem diagnosztizált személyek. Ahogy korábban beszámoltam róla, a Facebook minden felhasználót bizonyos szempontokat figyelembe véve egy-egy érdeklődési kategóriába sorol és ez alapján különböző hirdetéseket ajánl fel számukra. A kutatási kérdésem megvizsgálására a Facebook által létrehozott érdeklődési kategóriákat tartalmazó adatbázist használtam.

3.5. A függő és a kontroll változók létrehozása

A korábbi kutatások [\(5\)](#) [\(14\)](#) szerint a PHQ-9 kérdései mellett, hogy megfelelnek a DSM-IV diagnosztikai kritérium 9 elemének, megbízhatónak és érvényesnek tekinthetőek a depresszó súlyosságának értékeléséhez. Ebből adódóan a résztvevők kissé módosított PHQ-9 kérdésekre adott válaszait használtam fel a depresszió súlyosságának megítéléséhez. A módosított PHQ-9 kérdőívre adott válaszok pontszámait összesítettem és az eredeti pontozási kategóriák alapján arányosítottam. Mivel ebben az adatfelvételben 7 fokú skálát alkalmaztak, így az összpontszám 63 pont volt. Ennek megfelelően az új pontozási kategóriákat meghatároztam és a kérdőív válaszait tartalmazó adatbázisban eszerint kiszámoltam, hogy az egyes válaszadók melyik csoportba esnek.

Depresszió súlyossága	Pontszám	Részvevők száma
Nincs	0 – 10 pont	5 fő
Enyhe	11 – 21 pont	52 fő
Közepes	22 – 33 pont	63 fő
Közepesen súlyos	34 – 44 pont	22 fő
Súlyos	45 - 63 pont	8 fő

2. Táblázat Az arányosított pontozási rendszer alapján meghatározott depresszió súlyossági szintek

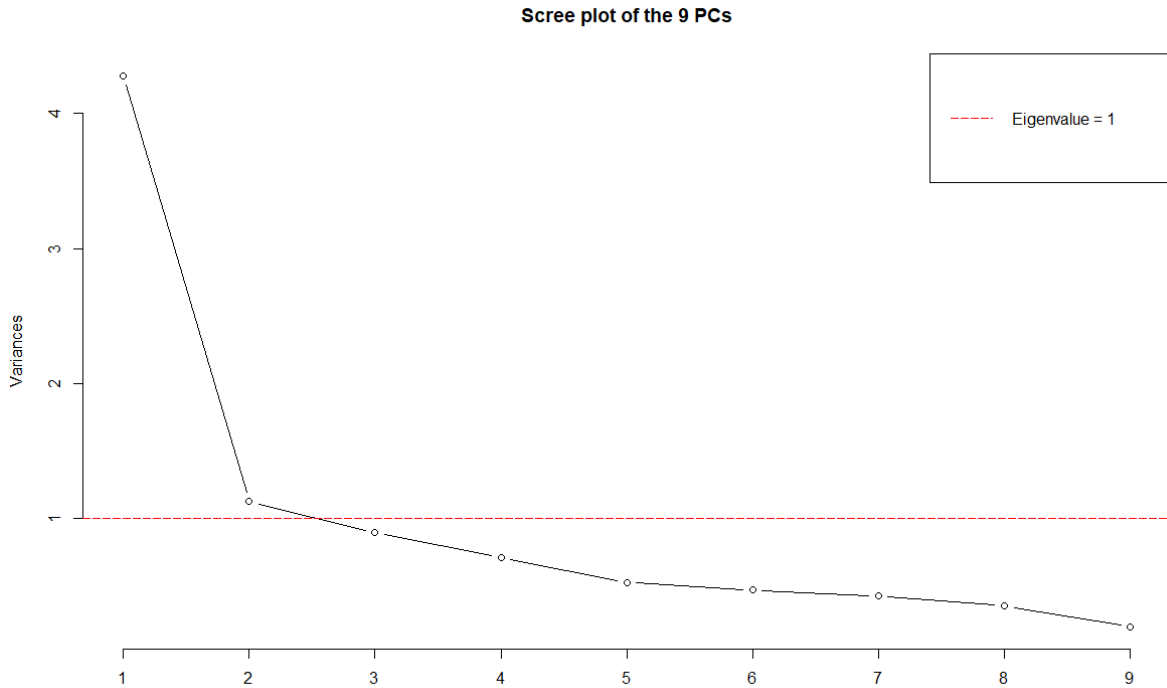
Emellett a korábbi tanulmányokban leírtak szerint meghatároztam egy küszöbértéket [\(6\)](#). Azokat a válaszadókat tekintetem depresszióban szenvedő személynek, akik 34 pont feletti értékkel rendelkeztek, azaz a közepesen súlyos vagy súlyos depressziós kategóriába tartoztak. A két csoport elkülönítésére egy 0-1 értékű dummy változót létrehoztam. A fent említett pontozás szerint a minta 80%-a, azaz 120 személy nincs depresszióval diagnosztizálva, míg a minta 20%-a depressziós anamnézissel rendelkezik. A nemenkénti megoszlás szerint 26 nő és 4 férfi szenved a módosított PHQ-9 pontozás szerint a mentális betegséggel, valamint 87 nő és 33 férfi nem szenved depresszióban.

Boothroyed és munkatársai tanulmányukban [\(14\)](#) arról számoltak be, hogy számos kutatásban vizsgálták a PHQ-9 megbízhatóságát és érvényességét, továbbá azt, hogy az egy vagy a két faktoros struktúra használata bizonyul jobbnak. Egyes kutatásokban egy faktoros modelleket használtak, míg más kutatók két faktoros modell mellett döntöttek. Egyik és másik modell mellett is számos érv és ellenérv szól, de hogy melyik modell bizonyul jobbnak, arra nincsen egyértelmű válasz. Néhány kutatás eredménye szerint csupán egy faktor rosszul illeszkedő modellt eredményez. Például Titov és munkatársai (2011) kutatásukban [\(23\)](#) 172 depressziós betegből álló mintát elemezve, azt az eredményt kapták, hogy az egy faktoros modell rosszul illeszkedik. Ezzel szemben Gonzalez-Blanch és munkatársai (2018) [\(11\)](#) szerint az egy és a két faktoros modell is jó illeszkedést ad, viszont mivel a két faktor szorosan összefügg egymással, így az egy faktoros modell előnyösebb takarékosági szempontból. Továbbá Beard és munkatársai (2016) [\(3\)](#) egy nagyobb minta (n=1023) vizsgálatánál szintén a két faktoros megoldás mellett döntöttek. Utóbbi kutatás esetében a megerősítő (konformatikus) faktoranalízis alkalmazásával született a két faktoros megoldás. A tanulmányban az első faktor az érdeklődés elvesztésére, depressziós hangulatra, értéktelenség érzetére, öngyilkos gondolatokra vonatkozó kérdéseket, vagyis a kognitív-affektív tüneteket tartalmazta. A második faktor az alvási nehézségekre, fáradtságra, étvágyzavarokra, koncentrációs nehézségekre, mozgás vagy beszéd problémákra utaló kérdéseket foglalta magába, tehát a szomatikus tüneteket tömörítette egybe.

Kognitív-affektív faktor	Szomatikus faktor
Kevés érdeklődés vagy örömmérség tevékenységei során	Nehezen tudott elaludni, éjszaka könnyen felébredt vagy túl sokat aludt
Szomorúság, lehangoltság vagy reménytelenség	Fáradtság vagy kevés energia
Rossz érzések saját magával kapcsolatban, vagy olyan gondolatok, hogy Ön sikertelen, vagy csalódást okozott önmaga vagy családja számára	Rossz étvágy vagy túlzott evés
Olyan gondolatok, hogy jobb lenne meghalni, vagy hogy valamilyen módon kárt tehetne önmagában	Koncentrációs nehézségek, például újságolvasás vagy tévénézés közben
	Mozgása vagy beszéde annyira lelassult, hogy mások is észrevehették, vagy ellenkezőleg, olyan nyugtalan volt, hogy a szokásosnál sokkal többet mozgott

3. Táblázat A feltáró faktoranalízis után kapott két faktor

A közösségi média viselkedési mintáinak nem csak a depresszió diagnózisával való kapcsolatát szerettem volna megvizsgálni, hanem az egyes depressziós skálákkal is. Ezáltal a depresszió diagnózisának változója mellett, a depresszió skála változókat is létehoztam. A módosított PHQ-9 kérdőív 9 kérdését Főkomponens elemzéssel megvizsgáltam, hogy az adathalmaz lehetséges egymással korreláló változóit lineárisan korrelálatlan változókká alakítsam. Ezzel az eljárással a dimenzió csökkentése mellett a lehető legnagyobb varianciát tudom megőrizni. A vizsgálat előtt ellenőriztem a 9 változó korrelációs mátrixának vizsgálatával a változók közötti korrelációt. A Kaiser-Meyer-Olkin teszt értéke alapján ($KMO > 0.8$) a változók közötti korreláció "dicséretes". Továbbá a Barlett tesztet is lefuttattam, ami azt vizsgálja, hogy a változószett korrelációs mátrixának főátlón kívüli elemei csak a véletlen miatt térnek el a nullától. Mivel az eredmény nem szignifikáns ($P < 0.05$), így a null hipotézist, miszerint a változók függetlenek egymástól elutasítjuk. A főkomponens elemzés futtatása után látható (1. ábra), hogy valószínűleg két komponens lesz, hiszen a harmadik komponens esetében a sajátérték 1 alá esik.



1. **Ábra** A komponensek számának meghatározása scree plot segítségével

Mivel előbbi eredményeim szerint két dimenzió lett, így Faktor analízist futtattam, hogy megnézzem az egyes változók hogyan korrelálnak egymással. Annak érdekében, hogy a faktorok egymással korrelálatlanok maradjanak Varimax rotálást alkalmaztam a Faktor analízis során.

```
Call:
factanal(x = ME2, factors = 2, scores = "regression", rotation = "varimax")

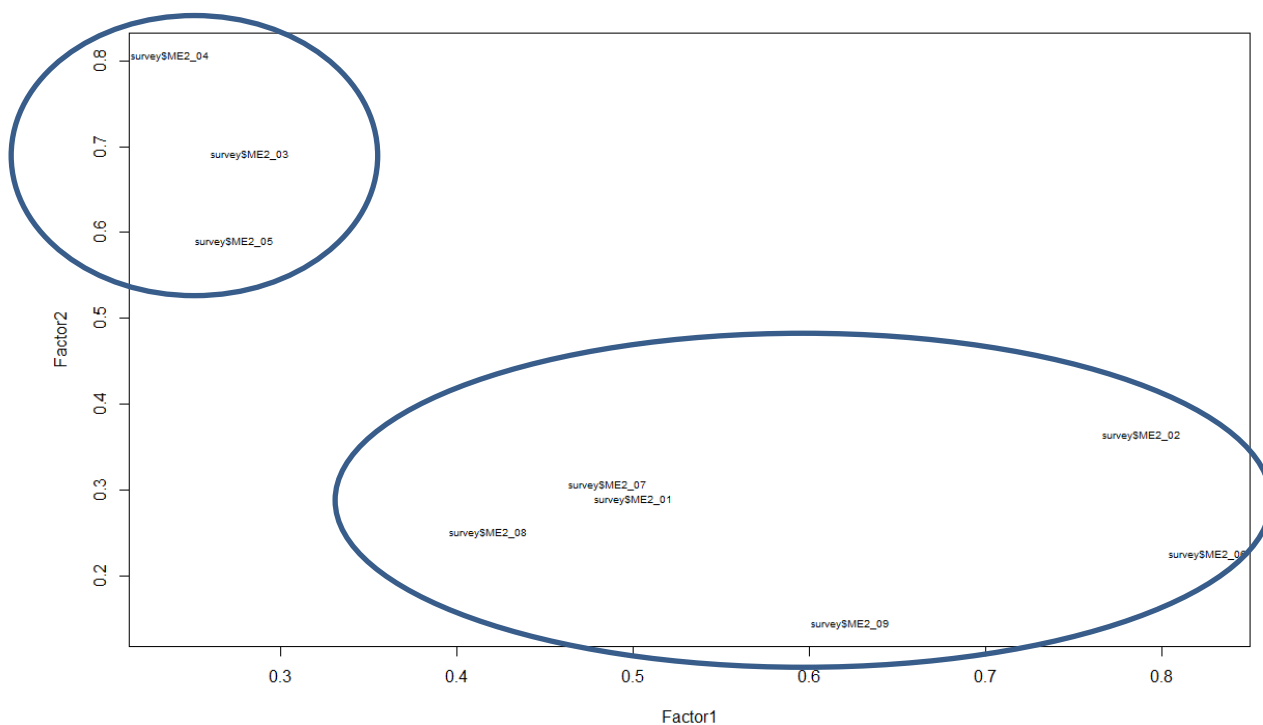
Uniquenesses:
survey$ME2_01 survey$ME2_02 survey$ME2_03 survey$ME2_04 survey$ME2_05 survey$ME2_06 survey$ME2_07
0.666          0.245          0.443          0.295          0.577          0.266          0.670
survey$ME2_08 survey$ME2_09
0.762          0.590

Loadings:
          Factor1 Factor2
survey$ME2_01 0.501  0.289
survey$ME2_02 0.789  0.364
survey$ME2_03 0.283  0.691
survey$ME2_04 0.237  0.805
survey$ME2_05 0.274  0.590
survey$ME2_06 0.827  0.225
survey$ME2_07 0.486  0.306
survey$ME2_08 0.418  0.251
survey$ME2_09 0.623  0.145

          Factor1 Factor2
SS loadings  2.568  1.919
Proportion Var 0.285  0.213
Cumulative Var 0.285  0.499

Test of the hypothesis that 2 factors are sufficient.
The chi square statistic is 47.76 on 19 degrees of freedom.
The p-value is 0.000278
```

2. **Ábra** Faktor analízis eredménye varimax rotálással



3. Ábra A PHQ-9 kérdései Faktor analízis vizsgálata után két faktort eredményeznek

Az első faktor az érdeklődés elvesztését, reménytelenséget, szomorúságot, önvádaskodást, koncentrációs nehézségeket, mozgás és beszéd lassulást, valamint a halálvágyat érintő módosított PHQ-9 kérdéseket foglalta magába. A második faktor az alvászavarra, fáradékonyságra, illetve az étvágyzavarra vonatkozó kérdéseket tömörítette magába (3. ábra). Ezáltal elmondható, hogy a korábbi szakirodalomban (14) leírt eredményekhez hasonlóan szintén két faktor lett, mégpedig a kognitív-affektív depresszió skála, illetve a szomatikus depresszió skála. Fontos megemlíteni, hogy Beard és munkatársai (2016) tanulmányától (3) eltérnek az egyes faktorokhoz tartozó kérdések. A vizsgált minta szerint az első faktor tömöríti magába a koncentrációs nehézségeket, illetve a mozgás és beszéd lassulást érintő kérdéseket, míg az említett tanulmányban azok a második faktorba kerültek. A faktorok segítségével már nem csak a depresszió diagnózisát tudtam a korábbi tanulmányokból ismert jellemzőkkel összevetni, hanem a különböző depressziós skálákkal való kapcsolatuk is vizsgálható volt.

A depresszió diagnózisa, illetve a depressziós skála változók és a viselkedési jellemzők közötti kapcsolat pontosabb felderítése érdekében célszerű kontroll változókat bevonni a modellekbe. A bemutatott szakirodalmak a regressziós modellek építése esetén demográfiai tényezők szerint kontrollálnak. A kérdőív válaszait tartalmazó adatbázis korra vonatkozó adataiból létrehoztam egy kétértékű változót, a 30 év alatti személyek esetében egyes értéket vett fel, a 30 éves vagy idősebb személyek esetében nullás értéket. A nem változó esetében átkódoltam a korábban meghatározott értékeket, így az egyes érték a férfiakra vonatkozott, a nullás érték a nőkre (korábban az egyes érték vonatkozott a férfiakra, a kettes érték a nőkre). A modellekbe a létrehozott kor és nem változót használtam kontroll

változóként. Első sorban a kontroll változók és a függő változók kapcsolatát vizsgáltam regressziós modellek segítségével.

A logisztikus regressziós modellbe bevontam a depresszió diagnózisát függő változóként, valamint a kor és a nem változót független változóként (4. ábra). Az eredményekből látható, hogy a kor 10%-os szinten szignifikáns kapcsolatot mutat a depresszióval. A nem változó értékét állandóként tekintve elmondható, hogy a 30 év alatti személyek esetében a depresszió bekövetkezésének log esélye 0,9688-cal nagyobb, vagyis $e^{0,9688}=2,63$ -szorosa a 30 éves vagy afelettekéhez képest.

```
Call:
glm(formula = depr_dummy ~ harminc_alatt + ferfi, family = "binomial",
     data = survey2)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.7916 -0.7916 -0.5483 -0.3701  2.3302

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.9687     0.5008  -3.931 8.44e-05 ***
harminc_alatt  0.9688     0.5351   1.811  0.0702 .
ferfi        -0.6778     0.5897  -1.149  0.2504
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 150.12  on 149  degrees of freedom
Residual deviance: 143.52  on 147  degrees of freedom
AIC: 149.52

Number of Fisher Scoring iterations: 4
```

4. Ábra A depresszió diagnózisa és a demográfiai változók közti kapcsolat vizsgálata logisztikus regresszióval

Ezt követően a kognitív-affektív és a szomatikus depresszió skálák és a demográfiai változók közti kapcsolatot is megvizsgáltam lineáris regresszióval (5. ábra). A kor változó 1%-os szinten szignifikáns korrelációt mutat a kognitív-affektív depresszió faktoral. A regressziós együtthatók vizsgálatából kiderül, hogy a kor változó egy egységnyi növekedése a kognitív-affektív depressziós faktor pontját 0,4866-tal növeli. Tehát a nem változót állandónak tekintve, a 30 év alatti személyek esetében 0,4866-tal magasabb a kognitív-affektív depressziós faktor pontja, mint a 30 éves vagy idősebb résztvevőknél.

```

Call:
lm(formula = kogn_aff ~ harminc_alatt + ferfi, data = survey2)

Residuals:
    Min       1Q   Median       3Q      Max
-1.6340 -0.6024 -0.2215  0.5438  2.5013

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.33363    0.14281  -2.336  0.02083 *
harminc_alatt  0.48661    0.15811   3.078  0.00249 **
ferfi         0.03742    0.17290   0.216  0.82895
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8835 on 147 degrees of freedom
Multiple R-squared:  0.06252, Adjusted R-squared:  0.04977
F-statistic: 4.902 on 2 and 147 DF, p-value: 0.008691

```

5. Ábra A kognitív-affektív skála és a demográfiai változók közti kapcsolat vizsgálata lineáris regresszióval

A szomatikus depresszió dimenziót függő változóként bevontam a lineáris regressziós modellbe (6. ábra). Ebben az esetben a kor változó mellett a nem változónak is 5%-os szignifikancia szinten szignifikáns hatása volt a szomatikus depresszió skálára. A regressziós együtthatók alapján a nem változót állandónak véve, a kor változó egy egységnyi növekedése a szomatikus depresszió faktor pontját 0,3508-cal növeli. Tehát a 30 év alattiak esetében a szomatikus depresszió faktor pontja 0,3508-cal magasabb a 30 éves vagy afelettekhez képest. Továbbá a kor változó hatásainak figyelembevétele után az is látható, hogy a férfiak esetében a szomatikus faktor pontja 0,3357-tel kevesebb a nőkhöz képest.

```

Call:
lm(formula = szomat ~ harminc_alatt + ferfi, data = survey2)

Residuals:
    Min       1Q   Median       3Q      Max
-1.93265 -0.59391 -0.01602  0.62242  2.11658

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.1511    0.1356  -1.114  0.2670
harminc_alatt  0.3508    0.1501   2.337  0.0208 *
ferfi        -0.3357    0.1642  -2.045  0.0426 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8388 on 147 degrees of freedom
Multiple R-squared:  0.08045, Adjusted R-squared:  0.06794
F-statistic: 6.431 on 2 and 147 DF, p-value: 0.002102

```

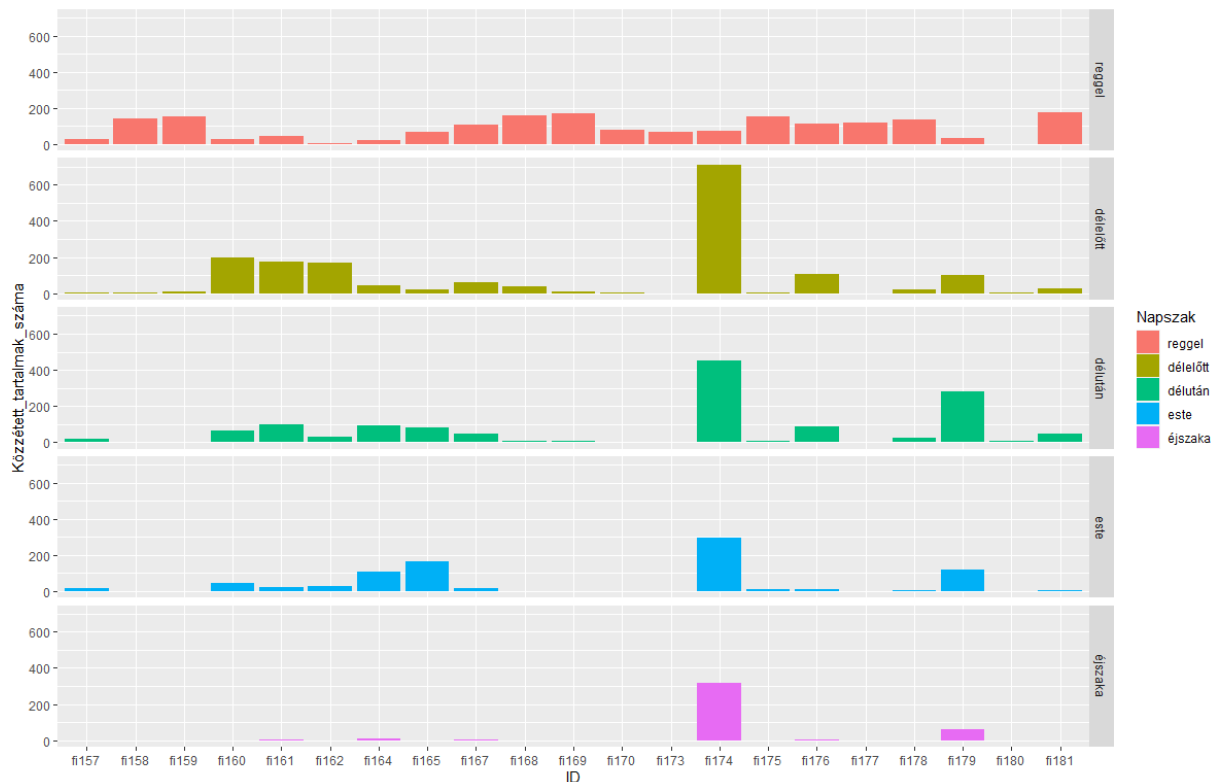
6. Ábra A szomatikus skála és a demográfiai változók közti kapcsolat vizsgálata lineáris regresszióval

3.6. Elemzés

3.6.1.1.1. *Az éjszakai fokozottabb aktivitás vizsgálata*

A rendelkezésre álló adatokon a korábbi kutatásokban bemutatott viselkedési mintázatokat teszteltem. Az első kutatási kérdés [\(K1\)](#) vizsgálatára, miszerint van-e eltérés a depressziós és nem depressziós emberek napszakonkénti aktivitásában, a felhasználók tevékenységeit tartalmazó adatbázis 2019 évi adatait használtam. Ahogy korábban említettem, az adatfelvétel 2019 áprilistól szeptemberig tartott. Ennél fogva az egyes felhasználóknak az adatbázisban csupán 2019 nyaráig vannak adatai, míg más felhasználóknak még a szeptemberi Facebook tevékenysége is le van kérve. Emiatt a teljes 2019 évi adatok helyett az első hat hónapból származó adatot kértem le, ami összesen 13.693 Facebook tevékenységet jelentett. A leszűrt adatbázisban felhasználónként látható napra, órára pontosan az általuk 2019 januártól júniusig közzétett tartalmak az említett 19 kategóriába sorolva. De Choudhury és munkatársai (2013) [\(6\)](#) eredményéből kiindulva a hipotézisem az volt, hogy az esti, valamint az éjszakai órákban a depresszióban szenvedő személyek nagyobb aktivitást mutatnak.

A korábbi kutatásban [\(5\)](#) az éjszakai aktivitás depresszióra gyakorolt hatását éjszakai és nappali időablak mutatók létrehozásával vizsgálták. Ezáltal az adatbázisban a tartalmakat öt napszakba soroltam a közzététel időpontja szerint. Ha a bejegyzések időpontja 5 és 7 óra közé esett, akkor reggeli napszakba soroltam. Amennyiben 8 és 12 óra közé, akkor délelőtti, 13 és 19 óra közé, akkor délutáni tartalomnak minősítettem. Továbbá, ha 20 és 22 óra között került ki a bejegyzés, akkor esti, ha 23 óra és reggel 4 óra között tették közzé, akkor éjszakai posztnak tekintettem. Ezután összesítettem felhasználókra napszakonként a posztok számát. Ezáltal látható volt, hogy az egyes felhasználók akár reggel vagy akár este mennyi tartalmat tettek közzé, ahogy a [7. ábra](#) is mutatja.



7. Ábra A 150 felhasználó közül 20 résztvevőnek a 2019.01. és 2019.06. közötti napszakonkénti posztjainak száma

Emellett az összes 2019 év első félévében közzétett tartalmak számát is kiszámoltam. Ebből minden felhasználó esetében megkaptam, hogy milyen aktivitási szokással rendelkeznek. Az aktivitási arány mutatók megmutatták, hogy milyen arányban oszlik meg a felhasználó posztolási szokása az öt napszakban. Ezt követően lekértem a meglévő adatokhoz a kérdőív adatait, hogy megvizsgáljam a depresszió diagnózisa és az egyes napszakok, valamint a depresszió skálák és az egyes napszakok közti korrelációt.

A létrehozott aktivitási arány mutatók és a depresszió diagnózisa kétértékű változó közti kapcsolatot logisztikus regressziós modellel vizsgáltam, kontrollálva a kor és a nem változókra. Hipotézisem szerint az esti, valamint az éjszakai órákban a depresszióban szenvedő személyek nagyobb aktivitást mutatnak. Ahogy a [8. ábráról](#) is leolvasható, a logisztikus regressziós modell futtatása után azt az eredményt kaptam, hogy 10%-os szignifikancia szinten nincsen szignifikáns kapcsolat a depresszió diagnózisa és az esti aktivitási arány között. Az éjszakai aktivitás vizsgálata során hasonló eredményt kaptam ([9. ábra](#)), tehát nincs szignifikáns korreláció az éjszakai aktivitás aránya és a depresszió diagnózisa között.

```

Call:
glm(formula = depr_dummy ~ est_ar + ferfi + harminc_alatt, family = "binomial",
    data = activ)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.1122  -0.7128  -0.5505  -0.3299   2.4664

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -2.2964    0.6094  -3.769 0.000164 ***
est_ar        2.1218    1.3296   1.596 0.110518
ferfi        -0.6962    0.6777  -1.027 0.304253
harminc_alatt 0.7907    0.5528   1.431 0.152569
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 136.43  on 137  degrees of freedom
Residual deviance: 128.35  on 134  degrees of freedom
(12 observations deleted due to missingness)
AIC: 136.35

Number of Fisher Scoring iterations: 5

```

8. Ábra A depresszió diagnózisa és az esti aktivitás közti kapcsolat vizsgálata logisztikus regresszióval

```

Call:
glm(formula = depr_dummy ~ éjsz_ar + ferfi + harminc_alatt, family = "binomial",
    data = activ)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.7818  -0.7795  -0.5530  -0.3655   2.3470

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.8006    0.5148  -3.497 0.00047 ***
éjsz_ar      -0.8141    5.7188  -0.142 0.88680
ferfi        -0.8666    0.6671  -1.299 0.19393
harminc_alatt 0.7717    0.5514   1.400 0.16166
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 136.43  on 137  degrees of freedom
Residual deviance: 130.84  on 134  degrees of freedom
(12 observations deleted due to missingness)
AIC: 138.84

Number of Fisher Scoring iterations: 5

```

9. Ábra A depresszió diagnózisa és az éjszakai aktivitás közti kapcsolat vizsgálata logisztikus regresszióval

Ezt követően a kognitív-affektív és a szomatikus depresszió dimenziók esetében is megvizsgáltam, hogy mutatkozik-e szignifikáns kapcsolat a különböző napszakok aktivitási arányával. A lineáris regressziós modellbe bevontam az adott napszak aktivitási mutatóját független változóként, valamint kontrolláltam a nemre és a korra. Ahogy a kognitív-affektív skála esetén, úgy a szomatikus depresszió skálánál sem volt látható szignifikáns kapcsolat bármely napszak aktivitásával.

```

Call:
lm(formula = kogn_aff ~ éjsz_ar + ferfi + harminc_alatt, data = activ)

Residuals:
    Min       1Q   Median       3Q      Max
-1.4692 -0.6058 -0.2190  0.4744  2.4879

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.29012    0.15365  -1.888  0.06116 .
éjsz_ar      -1.82442    1.95287  -0.934  0.35187
ferfi        -0.01912    0.18413  -0.104  0.91744
harminc_alatt 0.45647    0.16750   2.725  0.00729 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8812 on 134 degrees of freedom
(12 observations deleted due to missingness)
Multiple R-squared:  0.06113, Adjusted R-squared:  0.04011
F-statistic: 2.908 on 3 and 134 DF, p-value: 0.03701

```

10. Ábra A kognitív-affektív skála és az éjszakai aktivitás közti kapcsolat vizsgálata lineáris regresszióval

3.6.1.1.2. Az időbeli aktivitás változékonyságának vizsgálata a bejegyzéseken keresztül

A legtöbb tanulmány [\(5\)](#) [\(6\)](#) [\(7\)](#) beszámolt a depresszióban szenvedő személyek aktivitásának fokozottabb változásáról. Általában egy depresszívebb időszak alatt a közösségi oldalakon való jelenlétükben csökkenő tendencia figyelhető meg. Ennél fogva a 2018.01.01. és 2018.06.30. közötti adatokat összevetettem a 2019.01.01. és 2019.06.30. közötti adatokkal. Megvizsgáltam, hogy a két időszak között a bejegyzések közzétételével töltött napok száma valóban jelentősebb változást mutat a depresszióban szenvedő személyeknél, mint a depresszióban nem szenvedőknél. Fontos megjegyezni, hogy a tevékenységi adatokat tartalmazó adatbázisban a tevékenység típusa is szerepel, így ez alapján szűrni tudtam a bejegyzés közzétételével kapcsolatos tartalom típusokra (1. bejegyzés közzététele valakinek az idővonalára, 5. bejegyzés közzététele a saját idővonalára, 8. bejegyzés közzététele egy csoportban). Ezáltal az eredeti 346.352 posztot tartalmazó adatbázist leszűkítettem 113.413 bejegyzés típusú adatra.

Leszűrve a két fél éves időszakot látható volt, hogy 2018 első hat hónapjában összesen 5320 bejegyzést, míg 2019 első hat hónapjában 5584 bejegyzést tettek közzé. Majd megnéztem, hány olyan nap volt 2018-ban, amikor a felhasználók legalább egy bejegyzést közzétettek. Ezután ugyanezt a 2019 év első hat havi adatára is megvizsgáltam, végül a két adatbázist összefűztem. Ezáltal lett egy változó 2018 évre és egy változó 2019 évre vonatkozóan, ami a hat hónap alatt bejegyzések közzétételével töltött napok számát mutatta. Ezen változók segítségével létre tudtam hozni egy eltérés mutatót, ami a 2018 és a 2019 év közötti bejegyzéssel töltött napok száma közti különbséget mutatta. A 150 résztvevőből 16 személy esetében a két időszak közötti napok száma nem mutatott változást, ami két dologgal magyarázható. Az első lehetőség, hogy egyik időszakban sem volt olyan nap, amikor a résztvevők közzétettek volna legalább egy bejegyzést, így a változás mutató értéke nulla. A második indok az lehet, hogy a két időszak alatt pontosan ugyanannyi napot töltöttek egy vagy több bejegyzés létrehozásával. Utóbbi esetében ez négy

felhasználóra volt igaz. A változás mutató segítségével megfigyeltem, hogy a depressziós jelek azonosításához a Facebook aktivitás változékonysága megfelelő jellemzőnek bizonyul-e.

A depresszió diagnózisa kétértékű változót függő változóként, a változás mutatót független változóként bevontam a logisztikus regressziós modellbe. Annak érdekében, hogy a magyarázó változót kontroll alatt tartsam, a nem és a kor változót is bevettem a modellbe. Ahogy a [11. ábrán](#) látható, 10%-os szignifikancia szinten az aktivitás változékonyságának szignifikáns hatása van a depresszió diagnózisára (P-érték<0.1). Ha a 2018-as és a 2019-es év első felében a bejegyzéssel töltött napok száma közti különbség eggyel növekszik, akkor ez a diagnózisban a depresszió megállapításának esélyét (odds ratio) $e^{0,02613}=1,0265$ -tel növeli. Ami azt jelenti, ha egy felhasználó két fél éves napi aktivitása közt eggyel nagyobb változás mutatkozik, az ~2,6%-kal növeli annak esélyét, hogy depresszióval diagnosztizálják. A 90%-os konfidencia-intervallum az odds növekedése esetén: $e^{0,02613 \pm 1,645 \cdot 0,01586}=(e^{0,052}; e^{0,0000403})=(1,00004; 1,053)$

```
Call:
glm(formula = depr_dummy ~ kulonbseg + ferfi + harminc_alatt,
     family = "binomial", data = adat)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.1955 -0.7576 -0.5451 -0.3312  2.4716

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -2.11178    0.51612  -4.092 4.28e-05 ***
kulonbseg    0.02613    0.01586   1.647  0.0995 .
ferfi       -0.76357    0.60457  -1.263  0.2066
harminc_alatt 1.14111    0.55353   2.062  0.0393 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 150.12  on 149  degrees of freedom
Residual deviance: 140.92  on 146  degrees of freedom
AIC: 148.92

Number of Fisher Scoring iterations: 5
```

11. Ábra A depresszió diagnózisa és a posztolási aktivitás változása közti kapcsolat vizsgálata logisztikus regresszióval

A következőkben lineáris regresszióval a kognitív-affektív depresszió skála és a tevékenységi aktivitás változása közti korrelációt figyeltem meg ([12. ábra](#)). Majd a szomatikus depresszió skála és a posztolási aktivitás változása közti kapcsolatot vizsgáltam ([13. ábra](#)). Mindkét modellben a korábbiakhoz hasonlóan kontrolláltam a korra és a nemre. A bejegyzés létrehozásával töltött napok száma alapján mért aktivitási változás 10%-os szignifikancia szinten sem a kognitív-affektív, sem a szomatikus depresszióval nem mutatott kapcsolatot.


```

Call:
lm(formula = kogn_aff ~ kulonbseg + ferfi + harminc_alatt, data = adat)

Residuals:
    Min       1Q   Median       3Q      Max
-1.5764 -0.6228 -0.1960  0.5671  2.4238

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.366578   0.143679  -2.551   0.0118 *
kulonbseg    0.009138   0.005859   1.560   0.1210
ferfi        0.030230   0.172127   0.176   0.8608
harminc_alatt 0.533107   0.160144   3.329   0.0011 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8792 on 146 degrees of freedom
Multiple R-squared:  0.07789, Adjusted R-squared:  0.05894
F-statistic: 4.111 on 3 and 146 DF, p-value: 0.007811

```

12. Ábra A kognitív-affektív skála és a posztolási aktivitás változása közti kapcsolat vizsgálata lineáris regresszióval

```

Call:
lm(formula = szomat ~ kulonbseg + ferfi + harminc_alatt, data = adat)

Residuals:
    Min       1Q   Median       3Q      Max
-1.90535 -0.62207 -0.03762  0.62505  2.13499

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.166181   0.137278  -1.211   0.2280
kulonbseg    0.004190   0.005598   0.748   0.4554
ferfi       -0.338980   0.164458  -2.061   0.0411 *
harminc_alatt 0.372137   0.153009   2.432   0.0162 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.84 on 146 degrees of freedom
Multiple R-squared:  0.08397, Adjusted R-squared:  0.06515
F-statistic: 4.461 on 3 and 146 DF, p-value: 0.004976

```

13. Ábra A szomatikus skála és a posztolási aktivitás változása közti kapcsolat vizsgálata lineáris regresszióval

3.6.1.1.3. Az időbeli aktivitás változékonyságának vizsgálata a reakciókon keresztül

Ezután egy másik Facebook aktivitási lehetőséget vizsgáltam, mégpedig a felhasználók reakcióit. A reakció adatok esetében is a 2018 és 2019 év első hat havi adatát használtam fel arra, hogy megtudjam, hány olyan nap volt ebben az időszakban, amikor valamilyen közzétett tartalomra reagáltak a felhasználók. Ahogy korábban említettem, a reakciókat tartalmazó adatbázis több, mint 1.800.000 reakciót tartalmazott. Leszűkítve a két fél éves időszakra, továbbra is nagy adatmennyiséggel tudtam dolgozni, hiszen 2018 januártól júniusig a felhasználók 130.329-szer használták a reakció funkciót. 2019 első fél éves időszakában 147.058-szor reagáltak egy-egy közzétett tartalomra. Mindkét időszak alatt ez esetben is meghatároztam, hogy hány olyan nap volt, amikor a felhasználók legalább egyszer reagáltak valamilyen tartalomra. Ezt követően szintén létrehoztam egy mutatót, ami ebben az esetben a két fél éves időszak reakcióval töltött napok arányát mutatta. Az arány mutató segítségével ismét megvizsgáltam, hogy a depresszió diagnózisa, valamint a két dimenziója és az átlagos tevékenységi aktivitás között tapasztalható-e szignifikáns kapcsolat.

Először a reakcióval töltött napok aránya és a depresszió diagnózisa közti korrelációt vizsgáltam. A logisztikus regressziós modellbe a korábbiakhoz hasonlóan a kétértékű változót bevontam függő változóként, a havi arány mutatót független változóként, valamint kontrolláltam a demográfiai változókra. A hipotézis szerint a depresszióban szenvedő személyek aktivitása nagy mértékű változást mutat, amely általában csökkenő tendenciában nyilvánul meg. Ahogy a [14. ábrán](#) látható, a modellben 5%-os szignifikancia szinten a reakciókon keresztül vizsgált aktivitás változása szignifikáns korrelációt mutat a depresszióval. A 2018-as és 2019-es év első felében a reakció küldésével töltött napok arányának egy egységnyi növekedése, a depressziós diagnózis megállapításának esélyét (odds ratio) $e^{-1,7667} = 0,1709$ -cel csökkenti. Tehát, ha egy felhasználó aktivitási arányában egy egységgel nagyobb változás mutatkozik, az ~17%-kal csökkenti annak esélyét, hogy depresszióval diagnosztizálják. A 95%-os konfidencia-intervallum az odds növekedése esetén: $e^{-1,7667 \pm 1,96 * 0,8492} = (e^{-3,431}; e^{-0,102}) = (0,032; 0,903)$

```
Call:
glm(formula = depr_dummy ~ arany + ferfi + harminc_alatt, family = "binomial",
    data = data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.2554  -0.7754  -0.5328  -0.2370   2.8391

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -0.2507     0.9159  -0.274   0.7843
arany        -1.7667     0.8492  -2.080   0.0375 *
ferfi        -0.6970     0.6108  -1.141   0.2538
harminc_alatt  1.0344     0.5510   1.877   0.0604 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 148.77  on 146  degrees of freedom
Residual deviance: 135.52  on 143  degrees of freedom
(3 observations deleted due to missingness)
AIC: 143.52

Number of Fisher Scoring iterations: 7
```

14. Ábra A depresszió diagnózisa és a reakciókon keresztül megfigyelt aktivitás változása közti kapcsolat vizsgálata logisztikus regresszióval

Ezután megvizsgáltam lineáris regresszióval a depresszió dimenzióit is annak érdekében, hogy teszteljem megfigyelhető-e a korreláció a depresszió egyes dimenzióinál is. Elsőként a kognitív-affektív depresszió skálát vontam be függő változóként a lineáris regressziós modellbe, független változóként pedig a létrehozott arány mutatót használtam. Annak érdekében, hogy a független változót kontroll alatt tartsam, a nem és a kor változót is bevontam a modellbe. Ez esetben nem mutatkozott 10%-os szignifikancia szinten kapcsolat az aktivitás változékonyságával ([15. ábra](#)). Majd a szomatikus depressziós skálát vontam be függő változóként a lineáris regresszióba és ismét kontrolláltam a demográfiai változókra. Ebben a modellben sem figyelhető meg szignifikáns kapcsolat ([16. ábra](#)).

```

Call:
lm(formula = kogn_aff ~ arany + ferfi + harminc_alatt, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-1.5460 -0.5689 -0.2093  0.5133  2.5026

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.255326  0.154093  -1.657  0.0997 .
arany       -0.006037  0.045603  -0.132  0.8949
ferfi       -0.047634  0.175246  -0.272  0.7862
harminc_alatt 0.412425  0.160770   2.565  0.0113 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8769 on 143 degrees of freedom
(3 observations deleted due to missingness)
Multiple R-squared:  0.05074, Adjusted R-squared:  0.03082
F-statistic: 2.548 on 3 and 143 DF, p-value: 0.05829

```

15. Ábra A kognitív-affektív skála és a reakciókon keresztül megfigyelt aktivitás változása közti kapcsolat vizsgálata lineáris regresszióval

```

Call:
lm(formula = szomat ~ arany + ferfi + harminc_alatt, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-1.9483 -0.6197 -0.0084  0.6112  2.0550

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.10024  0.14566  -0.688  0.4924
arany       -0.02142  0.04311  -0.497  0.6200
ferfi       -0.31779  0.16565  -1.918  0.0570 .
harminc_alatt 0.33696  0.15197   2.217  0.0282 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8289 on 143 degrees of freedom
(3 observations deleted due to missingness)
Multiple R-squared:  0.07812, Adjusted R-squared:  0.05878
F-statistic: 4.039 on 3 and 143 DF, p-value: 0.008595

```

16. Ábra A szomatikus skála és a reakciókon keresztül megfigyelt aktivitás változása közti kapcsolat vizsgálata lineáris regresszióval

3.6.1.1.4. Az ismerősökkel való kapcsolat változása a tartalmakra adott reakciókon keresztül

A felhasználók ismerőseikkel való szociális kapcsolatának változását a különböző tartalmakra adott reakciókon keresztül vizsgáltam. A reakciókat tartalmazó adatbázis kiváló volt erre a célra, mivel amellet, hogy szerepelnek benne a résztvevők különféle reakciói, amelyek tartalom típusonként kategorizálva vannak, valamint időbélyeggel el vannak látva, még az is megtudható belőle, hogy a felhasználó egy ismerősének vagy egy csoportnak a közzétett tevékenységére reagált. Ahogy már említettem, az adatbázis alapvetően 1.801.430 reakciót tartalmazott, de a szociális kapcsolatok vizsgálatához számomra a lényeges visszajelzések azok voltak, amiket az ismerősöknek küldtek. Az ismerősök nevei anonimizált módon szerepelnek itt is, ahogy a korábban bemutatott ismerős listát tartalmazó adatbázisban. Az anonimizált nevek mind „@” karakterrel kezdődnek, így le tudtam szűkíteni az adatbázist csupán az ismerősöknek adott reakciókra. Az így kapott segéd adatbázis 832.278 reakciót tartalmazott. A társadalmi kapcsolatok változásához a 2018 év első hat havi reakcióit hasonlítottam össze a 2019 év első hat havi adataival. Ezáltal leszűrtem először a 2018.01. és 2018.06. között létrejött reakciókat, ami 66.139 reakciót jelentett. Majd

megszámláltam, hogy ebben a hat hónapban egy felhasználónak ismerősönként mennyi reakciója volt.

Vizsgálatomban a közeli ismerősökkel való kapcsolat változása volt lényeges. Ennek oka az volt, hogy valószínűleg akivel a Facebookon távolabbi ismerős, azzal a való életben sincs szorosabb kapcsolata. Ezáltal a kettőjük közötti kapcsolat megszűnése nem feltétlen a betegségnek tudható be, így az összes ismerőssel folytatott reakció vizsgálata torzítaná az eredményt. A közeli, szorosabb ismerettséget az adott reakciók számában mértem, úgy gondoltam, hogy akivel távolabbi kapcsolatot ápol, annak ritkábban reagál egy bejegyzésére. A közeli barátok tevékenységei a Facebook algoritmusának köszönhetően gyakrabban megjelennek a hírfolyamban, illetve a szoros kapcsolatból kifolyólag kíváncsibbak a felhasználók az ő tartalmaikra, ezáltal gyakrabban reagálnak is rájuk. A közelebbi barátokat elsőként úgy definiáltam, hogy a felhasználó azon ismerősei, akinek a hat hónap alatt legalább 20 reakciót adott a tartalmaira, tehát átlagban havi 3-4 reakciót küldött. Ebben az esetben a minta nagy mértékben lecsökkent, a 2019-es adatok vizsgálatánál 150 főből csupán 59 résztvevőnek volt ezeknek a feltételeknek megfelelően közeli ismerőse. A nagyobb mintaméret érdekében más módszert választottam.

Ezt követően a közeli ismerősöket úgy határoztam meg, hogy a felhasználók havonta legalább 5 reakciót adjanak az ismerősök közzétett tartalmaira. Ugyan ez havi szinten magasabb küszöbértéknek tűnik, mint az első módszernél a hat hónap alatt legalább 20 reakció feltétele, mivel az körülbelül havi szinten 3-4 reakciót jelentene. Mégis ebben az esetben több felhasználónak volt közeli ismerőse. Ennek oka az volt, hogy több olyan eset előfordult, amikor a felhasználó egy ismerősének a hat hónap alatt összességében 20-szor nem reagált valamilyen közzétett tartalmára, viszont egyes hónapban vagy hónapokban legalább öt reakciót küldött neki. Ebből adódóan az első módszer szerint az ismerős nem volt közeli ismerőse, viszont a második módszer szerint igen.

A második módszert alkalmazva, a 2018-as adatok esetében a kutatás résztvevői közül 103 főnek volt legalább 1 közeli ismerőse. Ezután a 2019-es adatoknál is megvizsgáltam a közeli ismerősök számát. A 2019.01. és 2019.06. közötti időszakról 70.268 ismerősnek küldött reakciót tartalmazott az adatbázis. Ugyanazon feltétellel meghatározva a közeli ismerősök számát, ebben az esetben 101 felhasználónak volt legalább 1 közeli ismerőse. Fontos megjegyezni, hogy az elemzés elején a ismerősöknek küldött reakciók leválogatásánál néhány csoportnak küldött visszajelzés is belekerült a szűrésbe, hiszen „@” karakterrel kezdődött a nevük. Ezeket a csoportokat ellenőriztem és úgy találtam, hogy kiestek a közeli ismerősöket tartalmazó segédadatbázisból, mivel valószínűleg a felhasználó havi szinten ötnél kevesebb reakciót küldött az adott csoportokba létrehozott tartalmakra. Tehát így már tényleg csak az ismerősöknek küldött reakciók szerepeltek az aggregált adatbázisban. Ezt követően a 2018 és a 2019 évi közeli ismerősök számának változását tartalmazó normalizált mutatót létrehoztam. A két fél év alatti közeli ismerősök számának különbségét vettem, majd elosztottam a felhasználó ismerős számával. A normalizáláshoz az ismerősöket tartalmazó meta adatbázist is beolvastam és hozzáfűztem az eddigi adatokhoz.

Ezt követően megvizsgáltam a depresszió diagnózisának, illetve a depresszió dimenzióinak kapcsolatát a közeli ismerősök számának változásával. Elsőként logisztikus regressziós modellbe bevontam függő változóként a depresszió diagnózisát mutató

kétértékű változót, magyarázó változóként pedig a közeli ismerősök számát tartalmazó normalizált különbség mutatót, valamint kontrolláltam a nemre és a korra. Ahogy a [17.ábra](#) is mutatja, a két változó között nincs 10%-os szinten szignifikáns korreláció.

```
Call:
glm(formula = depr_dummy ~ valtozas + ferfi + harminc_alatt,
     family = "binomial", data = data_final.s)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.7980 -0.7924 -0.5706 -0.3681  2.3356

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.9736    0.5020  -3.931 8.45e-05 ***
valtozas      1.9043   12.1368   0.157  0.8753
ferfi        -0.6841    0.5909  -1.158  0.2470
harminc_alatt 0.9779    0.5387   1.815  0.0695 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 150.12 on 149 degrees of freedom
Residual deviance: 143.49 on 146 degrees of freedom
AIC: 151.49

Number of Fisher Scoring iterations: 4
```

17. Ábra A depresszió diagnózisa és a közeli ismerősök számának változása közti kapcsolat vizsgálata logisztikus regresszióval

A következő modellekbe a létrehozott két depresszió skála és a közeli ismerősök számának változása közötti korrelációt vizsgáltam lineáris regresszióval, kontrollálva a nemre és a korra. A kognitív-affektív skála és a szomatikus depresszió skála esetén sem tapasztalható kapcsolat a közeli ismerősök számában mutatkozó változással 10%-os szignifikancia szinten.

```
Call:
lm(formula = kogn_aff ~ valtozas + ferfi + harminc_alatt, data = data_final.s)

Residuals:
    Min       1Q   Median       3Q      Max
-1.6318 -0.6037 -0.2218  0.5062  2.5009

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.33566    0.14346  -2.340  0.02065 *
valtozas      1.06750    3.93046   0.272  0.78632
ferfi         0.03468    0.17374   0.200  0.84207
harminc_alatt 0.49085    0.15937   3.080  0.00248 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8863 on 146 degrees of freedom
Multiple R-squared:  0.063,    Adjusted R-squared:  0.04374
F-statistic: 3.272 on 3 and 146 DF,  p-value: 0.02299
```

18. Ábra A kognitív-affektív skála és a közeli ismerősök számának változása közti kapcsolat vizsgálata lineáris regresszióval

```

Call:
lm(formula = szomat ~ valtozas + ferfi + harminc_alatt, data = data_final.s)

Residuals:
    Min       1Q   Median       3Q      Max
-1.93044 -0.59393 -0.02531  0.63715  2.12156

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.1532    0.1362  -1.125   0.2625
valtozas      1.1130    3.7314   0.298   0.7659
ferfi        -0.3385    0.1649  -2.052   0.0419 *
harminc_alatt 0.3552    0.1513   2.348   0.0202 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8414 on 146 degrees of freedom
Multiple R-squared:  0.08101, Adjusted R-squared:  0.06213
F-statistic:  4.29 on 3 and 146 DF, p-value: 0.006199

```

19. Ábra A szomatikus skála és a közeli ismerősök számának változása közti kapcsolat vizsgálata lineáris regresszióval

3.6.1.1.5. A gyakran említett témák vizsgálata

Több korábban bemutatott tanulmányban [\(6\)](#) is említik, hogy bizonyos témával, mint a munkával, családdal, vallással, spiritualitással összefüggő szavak kapcsolatba hozhatóak a depresszióval. Továbbá arról is több kutatás beszámolt [\(6\)](#) [\(8\)](#) [\(16\)](#), hogy a depresszióba szenvedő személyek gyakran beszélnek a közösségi oldalakon egészségi állapotukról, kezelésükről. Ennek tanulmányozására a Facebook által létrehozott érdeklődési kategóriákat tartalmazó adatbázist vizsgáltam. Az érdeklődési kategóriákat manuálisan csoportosítottam a fent említett témáknak megfelelő csoportokba (munka, társas kapcsolatok, vallás, spiritualitás, egészség). Elsőként leválogattam az általánosabb kategóriákat, vagyis a kis betűvel kezdődőket. Ezáltal az elsőre nagy adathalmaznak tűnő, több, mint 105.000 kategóriát tartalmazó adatbázis az összes felhasználóra nézve 25.704 kategóriára csökkent. Ez 944 egyedi érdeklődési kategóriát jelentett, amit egyesével megvizsgáltam és a munka, társas kapcsolat, vallás, spiritualitás, illetve egészség címkével ellátott csoportba soroltam. Ebből létrehoztam öt olyan változót, amely segítségével láttam, hogy a felhasználóknak például a munka csoportba hány érdeklődési kategóriát generált a Facebook. Tehát ezek a változók tükrözik, hogy az egyes témákban mennyire érdekeltek a Facebook szerint a felhasználók. A 150 résztvevőből 5 olyan felhasználó volt, akik egyik csoportból sem kaptak érdeklődési kategóriákat.

	ad_interest
1	aerob edzés
2	allergia
3	alternatív gyógyászat
4	alvás
5	ápolás
6	diabetikus étrend
7	diéta (táplálkozás)
8	diplomás ápoló
9	egészséges táplálkozás
10	egészségügy
11	fitnesz
12	fitneszterem
13	fitneszversenyek
14	fizioterápia
15	fogászat
16	fogorvos
17	gluténmentes étrend
18	gyógypedagógia
19	izomzat és fitnesz
20	jóga
21	kórház
22	massacre
23	nifty
24	orvos
25	orvosság
26	pilates
27	szénhidrátszegény étrend
28	terápia
29	testedzés
30	testépítés
31	testsúly
32	vegán táplálkozás
33	vegetarianizmus
34	vegetáriánus konyha
35	wellness (alternatív gyógyítás)
36	zöltség
37	zsír

20. Ábra Az egészség címkével ellátott érdeklődési kategóriák

Ezt követően logisztikus regressziós modellbe megvizsgáltam, hogy a depresszió diagnózisa az egyes csoportokkal kapcsolatba hozható-e. A modellben ismét kontroll alatt tartottam a nem és a kor változóval a független változót ([21. ábra](#)). A munka témakör esetén 10%-os szignifikancia szinten már korreláció mutatkozott a depresszió diagnózisával. A felhasználóknál a munka kategóriába tartozó érdeklődési kategóriák számának eggyel való növekedése a depresszió diagnózisának esélyét 0,9386-tal ($e^{-0,06335}$) csökkenti. Tehát akik a

Facebook szerint a munkával kapcsolatos témákban érdekeltek ~93,9%-kal kisebb eséllyel lesznek depresszióval diagnosztizálva.

```
Call:
glm(formula = depr_dummy ~ munka + ferfi + harminc_alatt, family = "binomial",
     data = ad_interest_full)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.0429 -0.7371 -0.5513 -0.3198  2.3492

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.23680    0.61476  -2.012  0.0442 *
munka        -0.06335    0.03315  -1.911  0.0560 .
ferfi        -0.68880    0.60280  -1.143  0.2532
harminc_alatt  0.91192    0.54628   1.669  0.0951 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 150.12  on 149  degrees of freedom
Residual deviance: 139.41  on 146  degrees of freedom
AIC: 147.41

Number of Fisher Scoring iterations: 5
```

21. Ábra A depresszió diagnózisa és a munka témájú érdeklődési kategória közti kapcsolat vizsgálata logisztikus regresszióval

A depresszió dimenzióit lineáris regresszióval vizsgáltam, miközben a demográfiai változók segítségével kontroll alatt tartottam a független változót. A kognitív-affektív depresszió skála esetén 10%-os szinten nem találtam szignifikáns kapcsolatot a munka kategóriával. A szomatikus depressziós skála és a munka téma között 1%-os szinten szignifikáns kapcsolat figyelhető meg (22. ábra). A standarizált együtthatókat megfigyelve látható, hogy a munka szórásának egységnyi változása 0,2415 szórás csökkenést eredményez a szomatikus depressziós skála esetében.

```
Call:
lm(formula = szomat ~ munka + ferfi + harminc_alatt, data = ad_interest_full)

Residuals:
    Min       1Q   Median       3Q      Max
-1.95560 -0.62428 -0.07599  0.63178  1.87791

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.182869    0.169541   1.079  0.28254
munka       -0.025547    0.008167  -3.128  0.00213 **
ferfi       -0.328764    0.159473  -2.062  0.04102 *
harminc_alatt  0.309230    0.146419   2.112  0.03639 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8148 on 146 degrees of freedom
Multiple R-squared:  0.1382,    Adjusted R-squared:  0.1205
F-statistic: 7.805 on 3 and 146 DF,  p-value: 7.217e-05
```

```
call:
lm(formula = szomat ~ munka + ferfi + harminc_alatt, data = ad_interest_full)

Standardized Coefficients:
(Intercept)           munka           ferfi  harminc_alatt
  0.0000000   -0.2414855   -0.1636649    0.1683441
```

22. Ábra A szomatikus skála és a munka témájú érdeklődési kategória közti kapcsolat vizsgálata lineáris regresszióval

A társas kapcsolatok vizsgálatánál sem a depresszió diagnózisa, sem a depresszió skálák esetén nem találtam szignifikáns kapcsolatot a társadalmi kapcsolat témájú érdeklődési kategóriákkal.

A vallást, mint érdeklődési témát szintén lineáris regressziós modellben vizsgáltam a szomatikus depressziós skálával, kontrollálva a nemre és a korra. Ahogy a [23. ábrán](#) látható, 5%-os szinten szignifikáns korreláció figyelhető meg. A standarizált együtthatókból látható, hogy a vallás szórásának egységnyi változása 0,16797 szórás csökkenést eredményez a szomatikus depresszió skála esetében. A kognitív-affektív depresszió skálával, valamint a depresszió diagnózisával nem mutatott 10%-os szinten szignifikáns korrelációt a vallás téma.

```
Call:
lm(formula = szomat ~ vallas + ferfi + harminc_alatt, data = ad_interest_full)

Residuals:
    Min       1Q   Median       3Q      Max
-1.97698 -0.61162 -0.04429  0.72338  2.00321

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.0001321  0.1516910  -0.001  0.9993
vallas      -0.1235571  0.0582036  -2.123  0.0355 *
ferfi       -0.3732554  0.1631920  -2.287  0.0236 *
harminc_alatt 0.3075759  0.1497407   2.054  0.0418 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.829 on 146 degrees of freedom
Multiple R-squared:  0.108,    Adjusted R-squared:  0.08966
F-statistic: 5.892 on 3 and 146 DF,  p-value: 0.0007991
```

```
Call:
lm(formula = szomat ~ vallas + ferfi + harminc_alatt, data = ad_interest_full)

Standardized Coefficients:
            (Intercept)          vallas          ferfi harminc_alatt
            0.0000000      -0.1679745      -0.1858134      0.1674434
```

23. Ábra A szomatikus skála és a vallás témájú érdeklődési kategória közti kapcsolat vizsgálata lineáris regresszióval

A spiritualitás érdeklődési kategória esetén nem találtam szignifikáns kapcsolatot 10% szignifikancia szinten a depresszió diagnózisával vagy valamely depressziós dimenzióval.

Végül az egészséggel, kezelésekkel kapcsolatos érdeklődési kategóriákat tömörítő csoport esetén az előzőekhez hasonlóan logisztikus regressziós modellben vizsgáltam a depresszió diagnózisával való kapcsolatot, lineáris regresszióval a depresszió skálákat, miközben kontrolláltam a demográfiai tényezőkre. A három modell eredménye alapján csupán a kognitív-affektív depresszió skála és az egészség érdeklődési kategória között volt szignifikáns kapcsolat 10%-os szinten. A standarizált együtthatókból leolvasható, hogy az egészség szórásának egységnyi változása 0,1775 szórás csökkenést eredményez a kognitív depresszió skála esetében.

```

Call:
lm(formula = kogn_aff ~ egeszseg + ferfi + harminc_alatt, data = ad_interest_full)

Residuals:
    Min       1Q   Median       3Q      Max
-1.5556 -0.6191 -0.1656  0.5188  2.4526

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.07443    0.18728  -0.397  0.69163
egeszseg    -0.02847    0.01352  -2.106  0.03690 *
ferfi       -0.08965    0.18125  -0.495  0.62163
harminc_alatt 0.41838    0.15961   2.621  0.00969 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8733 on 146 degrees of freedom
Multiple R-squared:  0.09017, Adjusted R-squared:  0.07147
F-statistic: 4.823 on 3 and 146 DF, p-value: 0.003127

```

```

Call:
lm(formula = kogn_aff ~ egeszseg + ferfi + harminc_alatt, data = ad_interest_full)

Standardized Coefficients::
(Intercept)          egeszseg          ferfi harminc_alatt
0.000000000  -0.17749084   -0.04278084    0.21833854

```

24. Ábra A kognitív skála és az egészség témájú érdeklődési kategória közti kapcsolat vizsgálata lineáris regresszióval

4. Az eredmények értékelése

A viselkedési minták tesztelésénél elsőként a felhasználók napszakonkénti aktivitási szokásait figyeltem meg. Az ehhez létrehozott aktivitási arány mutatók segítségével, mind az öt napszak aktivitási arányát össze tudtam vetni a depresszió diagnózisával, illetve a két depresszió skálával. A reggeli, délelőtti, délutáni napszakok esetében szignifikáns korreláció sem a depresszió diagnózisával, sem a kognitív-affektív vagy a szomatikus depresszió skálával nem mutatkozott. A hipotézisem szempontjából az izgalmasabb kérdés az volt, hogy az esti és az éjszakai aktivitás milyen kapcsolatban áll a három változóval. A depresszió diagnózis vizsgálata során a logisztikus regressziós modell eredményei szerint nincs szignifikáns kapcsolat sem az esti, sem az éjszakai aktivitással. A lineáris regressziós modellel vizsgált depresszió skálák esetében ugyancsak ez mondható el. Tehát a depresszióval azonosított személyeknél az esti, éjszakai fokozottabb közösségi média jelenlét nem figyelhető meg a vizsgált mintában, ezáltal hipotézisem elutasítom. Az első kutatási kérdésemet [\(K1\)](#) megválaszolva, a depressziós és nem depressziós emberek napszakonkénti aktivitásában nem mutatkozott eltérés. Fontos itt újra megjegyezni, hogy Eichstaedt és munkatársai (2018) tanulmányukban [\(8\)](#) szintén arra a megállapításra jutottak, hogy nincsen a két csoport időbeli tevékenységi szokásaiban szignifikáns különbség.

Második tesztelt viselkedési minta az aktivitás változása volt, amit két időszak összehasonlításával figyeltem meg. Az aktivitást két szempontból is vizsgáltam. Elsőként a bejegyzéseken keresztül, majd a reakciókon keresztül. A bejegyzések vizsgálata a saját idővonalukra közzétett posztoknál szélesebb körre terjedt ki. Kutatásomba bevontam az ismerősök idővonalára és a csoportokba posztolt bejegyzéseket is. Az elemzés során a középpontban első sorban nem a bejegyzések vagy a reakciók mennyisége állt, inkább a bejegyzéssel vagy reakcióval töltött napok számára fókuszáltam. A napok számának kiszámításával mindkét aktivitás esetén létre tudtam hozni egy-egy mutatót. Ennek segítségével meg tudtam vizsgálni, hogy van-e nagyobb aktivitási változékonyság a depresszióban szenvedő személyeknél a depresszióban nem szenvedőkhöz képest [\(K2\)](#). A bejegyzés adatokon keresztül létrehozott mutató elemzéséből kiderült, hogy a depresszió diagnózisával 10%-os szinten szignifikáns kapcsolat volt tapasztalható, tehát az aktivitás változásában mutatkozó növekedés a depresszió diagnózisának esélyét növeli, de fontos megjegyezni, hogy a P-érték alig kevesebb, mint 10%, továbbá a depressziós dimenziókkal 10%-os szinten nem mutatott szignifikáns korrelációt. Ezáltal összességében egyértelműen kimutatható hatásról nem beszélhetünk a bejegyzés adatokon keresztül megvizsgált aktivitási változékonyság esetében.

A reakció adatok alapján létrehozott aktivitási arány mutató a depresszió dimenzióival szintén nem mutatott kapcsolatot 10%-os szignifikancia szinten, viszont a depresszió diagnózisával tapasztalható volt már 5%-os szinten szignifikáns korreláció. Az elemzésből kiderült, hogy az aktivitás arányának növekedésével a depresszió diagnózisának esélye csökken. Az aktivitási arány mutató a 2018/ 2019 évi reakcióval töltött napok számának

arányát mutatta, tehát aki a reakciókon keresztül mérve 2018-hoz képest 2019-ben aktívabb, azoknál nagyobb az esélye a depresszió diagnózisának.

A következő depressziós személyeket érintő viselkedési jellemző a társadalmi kapcsolatok változása volt. Ebben az esetben az ismerősöknek adott reakciókat vizsgáltam. A 2018 és 2019 év január és június között létrehozott reakciókból kiszámítottam, hogy a 150 felhasználónak mennyi közeli ismerőse van, valamint ez a szám a két időszak alatt mennyivel változott. A normalizált mutató segítségével megfigyeltem, hogy változnak-e a depresszióban szenvedő személyek társadalmi kapcsolatai [\(K3\)](#). A depresszió diagnózisa és a depresszió skálák vizsgálata esetén sem tapasztaltam 10%-os szinten szignifikáns hatást. Ezáltal arra jutottam, hogy ez a viselkedési jellemző nem bizonyul jó prediktornak a szociális visszahúzódas, mint depressziós tünet azonosításában.

Végezetül azokat a témákat vizsgáltam, amiről a szakirodalom szerint a depresszióban szenvedő személyek gyakran beszélnek a közösségi oldalaikon. A Facebook által generált érdeklődési kategóriákból létrehoztam öt csoportot, amely megmutatta, hogy a felhasználók mennyire érdekeltek a munka, a társas kapcsolatok, a vallás, a spiritualitás, valamint az egészség témákban, vagyis hány olyan érdeklődési kategóriát generált nekik a Facebook, ami ezen öt témakör valamelyikébe sorolható. Ezen öt mutatón keresztül vizsgáltam, hogy gyakrabban említik-e a depressziós emberek ezeket a témákat, mint a depresszióval nem diagnosztizált személyek [\(K4\)](#).

Elsőként a munka témakörnek a depresszió diagnózisával, valamint a depresszió skálákkal való kapcsolatát vizsgáltam. A depresszió diagnózisa és a szomatikus depresszió skála esetén 10%-os szinten szignifikáns korreláció volt. Akik a munka csoportból több érdeklődési kategóriát kaptak, azok kisebb eséllyel lesznek depresszióval diagnosztizálva. Továbbá akiknek a szomatikus depressziós skálán alacsony pontszámuk volt, nagyobb valószínűséggel kerülnek a munka címkével ellátott érdeklődési kategóriába. Tehát a munka témába érdekelt személyekre kevésbé jellemző a szomatikus depresszió.

A társadalmi kapcsolat, valamint a spiritualitás esetén nem volt tapasztalható 10%-os szignifikancia szinten kapcsolat.

A vallás téma esetén a szomatikus depressziós skálával már 5%-os szinten is megfigyelhető volt a szignifikáns kapcsolat, míg a másik két vizsgált változónál nem figyelhető meg még 10%-os szinten sem szignifikáns korreláció. Akiknek a szomatikus skálán alacsony pontszámuk volt, nagyobb valószínűséggel kerülnek a Facebook algoritmus által a vallás témájú érdeklődési kategóriákba. Tehát a vallás témába érdekelt személyekre kevésbé jellemző a szomatikus depresszió.

Az egészség téma tesztelésénél a szomatikus depressziós skála és a depresszió diagnózisa esetén 10%-os szinten nem volt megfigyelhető a szignifikáns kapcsolat. A kognitív-affektív dimenzió esetén már 5%-os szinten is látható volt a szignifikáns korreláció. Akiknek ezen a skálán alacsony pontszámuk volt, nagyobb valószínűséggel kerültek az

egészség témájú érdeklődési kategóriába. Tehát az egészség témába érdekelt személyekre kevésbé jellemző a kognitív-affektív depresszió.

5. Összefoglalás

A depresszió napjainkban egy világszintű betegség, amelyet a leggyakoribb mentális betegségek között tartanak számon. A betegségnek számos ismert tünete van, amik az érintett személyek mellett, a környezetük számára is felismerhetőek. Ennek ellenére a depresszióban szenvedő emberek sok esetben csak későn vagy sohasem kapják meg a megfelelő kezelést a gyógyuláshoz. A kezelés hiányában a tünetek súlyosabbá válhatnak, akár az öngyilkossági kísérletekig vezethetnek.

Napjainkban a közösségi oldalak rendkívül elterjedtek, a felhasználók napi szinten rengeteg tartalmat tesznek közzé saját és ismerőseik oldalán, vagy egy csoportban. A rögzített nagy mennyiségű adatokon keresztül megismerhetjük a felhasználók nyelvi és viselkedési szokásait, ezáltal információt kaphatunk szociális kapcsolataikról, valamint mentális állapotukról. Számtalan kutatásban a közösségi média ezen előnyét kihasználva, a felhasználók adatait elemezték a depressziós tünetek korai felismerése érdekében. A korábbi tanulmányokban a felhasználók nyelvi és viselkedési szokásait vizsgálták a depresszióban szenvedő és nem szenvedő személyek közti eltérő jellemzők megfigyeléséhez. Egyes tanulmányokban a felhasználók közösségi média adatai mellett egyéb szűrési eszközöket is alkalmaztak. A dolgozatban bemutatott kutatások a kínai és az angol felhasználók nyelvi és viselkedési jellemzőit vizsgálták. A szakirodalmi összefoglaló részben látható volt, hogy az eltérő kulturális és nyelvi ismérvek miatt problémát okozhat a nyelvi jellemzők vizsgálatából származó egyes kutatási eredmények összehasonlítása. A kutatók célja a depressziós és nem depressziós személyeket osztályozó modell létrehozása volt, de a speciális, kis számú minta, valamint a már említett eltérő kulturális és nyelvi jellemzők következtében az általánosíthatóság kérdéses.

Szaktervezésem célja a depresszió jeleinek azonosítása volt a közösségi média adatain keresztül, valamint a felhasználók által kitöltött online kérdőív segítségével. A fent említett általánosíthatósági problémából adódóan, kutatásom során a viselkedési mintázatokra fókuszáltam, a nyelvi jellemzők vizsgálatától eltekintettem. A kutatásban 2019-ben készült adatfelvételtől származó adatbázisokat használtam, amely 150 felhasználó Facebook archívumának adatait tartalmazta. A nagy mennyiségű adatok tartalmazták a felhasználók ismerőseinek a listáját, a tartalmakra adott reakcióikat, posztjaikat, kedvelt oldalaink listáját, valamint hirdetés céljára a Facebook által létrehozott érdeklődési kategóriák listáját. Továbbá a felhasználók által kitöltött kérdőív módosított PHQ-9 kérdéseire adott válaszok segítségével a depresszió súlyosságát is meg tudtam határozni, valamint a depresszió két dimenzióját is létrehoztam, a kognitív-affektív és a szomatikus depresszió faktort.

Egyes szakirodalmi eredmények alapján [\(6\)](#) a depressziós személyek az esti és az éjszakai órákban aktívabbak az alvási problémáiknak köszönhetően, ezáltal az első vizsgált viselkedési jellemző a napszakonkénti aktivitás volt. A napszakonként létrehozott aktivitási arány mutató segítségével információt kaptam a felhasználó posztolási szokásáról, valamint regressziós modell segítségével meg tudtam vizsgálni az adott napszak aktivitási aránya és a depresszió diagnózisa, valamint az egyes depresszió faktorok közti kapcsolatot. Az eredményeket látva, a depressziós és a nem depressziós személyek napszakonkénti aktivitásában nem volt eltérés.

Továbbá a szakirodalmi eredmények [\(5\)](#) [\(6\)](#) [\(7\)](#) arról számoltak be, hogy a depresszióban szenvedő embereknél fokozottabb aktivitási változékonyság figyelhető meg. Ezt két Facebook tevékenységgel töltött időszak összehasonlításával teszteltem. A felhasználók reakcióval és bejegyzéssel töltött napjait számláltam meg a változás mutató létrehozásához. A bejegyzéseken keresztül vizsgált aktivitási változékonyság esetében egyértelműen kimutatható hatás sem a depresszió diagnózisával, sem a kognitív-affektív vagy a szomatikus depresszió faktorra nem mutatkozott. A reakciók elemzése esetén kiderült, hogy az aktivitás arányának növekedésével a depresszió pozitív diagnózisának esélye csökken, vagyis akik 2018-hoz képest 2019-ben több reakciót küldtek ismerőseiknek, azoknál nagyobb az esélye a depresszió diagnózisának. A két faktor vizsgálatánál szignifikáns kapcsolat nem volt tapasztalható.

A korábbi kutatási eredmények [\(5\)](#) [\(6\)](#) [\(7\)](#) szerint a depressziós emberek szociálisan visszahúzódtak, leredukálódnak társas kapcsolataik. Ennek vizsgálatára szintén a felhasználók reakcióit tartalmazó adatokon keresztül hoztam létre a közeli ismerősök számának változását leíró mutatót. Ennek segítségével vizsgáltam, hogy a közeli ismerősök számában megfigyelhető változás kapcsolatot mutat-e a depresszió diagnózisával, vagy az egyes depresszió faktorokkal. A regressziós modellek futtatása után azt az eredményt kaptam, hogy nincsen egyik vizsgált változóval sem szignifikáns kapcsolata.

Korábbi vizsgálati eredményekre alapozva [\(6\)](#) [\(7\)](#) [\(8\)](#) [\(16\)](#) az utolsó vizsgált viselkedési jellemző a depressziós személyek által gyakran említett témák voltak, mint a munka, társas kapcsolatok, vallás, spiritualitás, valamint az egészség. Az eddigi kutatásokban ezt a szöveges adatok nyelvi jellemzőinek vizsgálatával tesztelték. A már említett nyelvi sajátosságokból eredő problémák elkerülése érdekében nem a szöveges adatok nyelvi jellemzőit vizsgáltam. Annak tesztelésére, hogy valóban gyakrabban beszélnek ezekről a témákról, a Facebook által generált érdeklődési kategóriákat használtam. Azt, hogy egy-egy érdeklődési kategória milyen témának, csoportnak felel meg a fent említett öt csoport közül, ezt magam határoztam meg. Majd létrehoztam egy mutatót, hogy mennyire érdekelt a felhasználó egy témában, tehát hány érdeklődési kategóriát hozott létre neki a Facebook. Megvizsgálva az öt téma alapján létrehozott mutatót a depresszió diagnózisával és dimenzióival, azt az eredményt kaptam, hogy a társas kapcsolatok, valamint spiritualitás esetén nem volt tapasztalható szignifikáns kapcsolat. A további témák vizsgálata esetén látható volt, hogy akik a munka csoportból több érdeklődési kategóriát kaptak, azok kisebb eséllyel lesznek depresszióval diagnosztizálva, valamint ebbe a témába érdekelt személyekre kevésbé jellemző a szomatikus depresszió. A vallás témába érdekeltre ez szintén igaz volt, tehát rájuk is kevésbé lesz jellemző a szomatikus depresszió. Míg az egészség témába érdekelt személyekre a kognitív-affektív depresszió lesz kevésbé jellemző.

A korábbi szakirodalmakban leírtakat több eredményem is alátámasztotta, az aktivitás növekedése vagy egyes érdeklődési témák hasonlóképp jó prediktorok voltak, ennek ellenére az eredményeket óvatosan kell kezelni. Szakdolgozatomban a kis számú, nem valószínűségi minta miatt, valamint az egyes demográfiai csoportoknál megfigyelhető eltérő Facebook használat miatt, a fent említett eredmények esetében az általánosíthatóság kérdésesnek bizonyul. A megbízhatóbb eredmények érdekében a regressziós modellek esetében kontroll változókat használtam, a nemre és a korra kontrolláltam. A kutatásom célja nem egy osztályozó modell felépítése volt, ami magas pontossággal diagnosztizálja a

résztevőket. Céлом inkább az volt, hogy megmutassam, hogy a Facebook által tárolt adatok egyfajta kiegészítő eszközként funkcionálhatnak a depressziós jelek azonosításában.

Végezetül bízom abban, hogy kutatásomban sikerült rávilágítani a közösségi média szerepére, miszerint a Facebook és más közösségi oldalak által tárolt adatok megfelelőek lehetnek egyéb diagnosztikai eszközök kiegészítőjeként. Ezáltal segíthetnek a mentális betegséggel küzdő személyek problémáinak korai felismerésében. Például a közösségi oldalak algoritmusába beépített figyelmeztető rendszeren keresztül, amely, ha egy felhasználónál több depressziós tünetre utaló jellemzőt érzékel, akkor szakemberek elérhetőségével kapcsolatos híreket vagy támogató csoportokat jelenít meg az érintett hírfolyamában, vagy ismerőseinek hívja fel a figyelmét egy barátjuk problémájára, ezzel támogatva a felhasználókat.

Irodalomjegyzék:

1. Alan M. Gruenberg, Reed D. Goldstein & Harold A. Pincus (2005). Classification of Depression: Research and Diagnostic Criteria: DSM-IV and ICD-10, *Biology of Depression. From Novel Insights to Therapeutic Strategies*. doi: <https://doi.org/10.1002/9783527619672.ch1>
2. Bagroy, S., Kumaraguru, P., & De Choudhury, M. (2017). A social media based index of mental well-being in college campuses. In *Proceedings of the 2017 CHI Conference on Human Factors in Computer Systems*. doi: <https://doi.org/10.1145/3025453.3025909>
3. Beard, C., Hsu, K. J., Rifkin, L. S., Busch, A. B., & Bjorgvinsson, T. (2016). Validation of the PHQ-9 in a psychiatric sample. *J. Affect. Disord.* 193, 267-273. doi:10.1016/j.jad.2015.12.075
4. Cheng, Q., Li, T. M., Kwok, C. L., Zhu, T., & Yip, P. S. (2017). Assessing Suicide Risk and Emotional Distress in Chinese Social Media: A Text Mining and Machine Learning Study, *Journal of Medical Internet Research*, 19(7), e243.
5. De Choudhury, M., Counts, S., Horvitz, E. J., & Hoff, A. (2014, February). Characterizing and predicting postpartum depression from shared facebook data. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing* (pp. 626-638).
6. De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013, June). Predicting depression via social media. In *Seventh international AAAI conference on weblogs and social media*.
7. De Choudhury, M., Kiciman, E., Dredze, M., Coppersmith, G., & Kuman, M. (2016). Discovering shifts to suicidal ideation from mental health content in social media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computer Systems* (pp. 2098-2110).
8. Eichstaedt, Johannes C., et al. "Facebook language predicts depression in medical records." *Proceedings of the National Academy of Sciences* 115.44 (2018): 11203-11208.
9. Eurostat (2018): Mental health and related issues statistics, https://ec.europa.eu/eurostat/statistics-explained/index.php/Mental_health_and_related_issues_statistics#Extent_of_depressive_disorders (Letöltés dátuma: 2020.09.14.)
10. Eurostat (2018): World Mental Health Day: data on chronic depression, <https://ec.europa.eu/eurostat/web/products-eurostat-news/-/EDN-20181010-1> (Letöltés dátuma: 2020.09.14.)
11. González-Blanch, C., Medrano, L. A., Muñoz-Navarro, R., Ruíz-Rodríguez, P., Moriana, J. A., Limonero, J.T., et al. (2018). Factor structure and measurement invariance across various demographic groups and over time for the PHQ-9 in primary care patients in Spain. *PloS one*, 13(2), e0193356.

12. Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2017). Detecting depression and mental illness on social media: an integrative review. *Current Opinion in Behavioral Sciences*, 18, 43-49.
13. Johannes Breuer, Zoltán Kmetty, Mario Haim, & Sebastian Stier: User-focused approaches for collecting Facebook data in the “post-API age”: Experiences from two studies and recommendations for future research, Kézirat.
14. Luke Boothroyd, David Dagnan, & Steven Muncer (2019). PHQ-9: One factor or two?. *Psychiatry Research*, 271, 532-534. doi: <https://doi.org/10.1016/j.psychres.2018.12.048>
15. Markos-Kujbus, É., & Gáti, M. (2012). A közösségi média mint online stratégiai eszköz, *Online marketing szekció*, 8.sz. tanulmány.
16. Moreno, M. A., et al. "Feeling bad on Facebook: Depression disclosures by college students on a social networking site." *Depression and anxiety* 28.6 (2011): 447-455.
17. Murray, C. J., & Lopez, A. D. (1996). *The Global Burden of Disease*. World Health Organization, Geneva. doi: <https://apps.who.int/iris/handle/10665/41864>
18. O'Connor, E. A., Whitlock, E. P., Beil, T. L., & Gaynes, B. N. (2009). Screening for depression in adult patients in primary care settings: A systematic evidence review. *Ann Intern Med* 151: 793–803.
19. Reece, A. G., Reagan, A. J., Lix, K. L. M., Dodds, P. S., Danforth, C. M., & Langer, E. J. (2016). Forecasting the Onset and Course of Mental Illness with Twitter Data, arXiv: 1608.07740.
20. Rihmer, Z., & Perczel Forintos, D. (2014): A depresszió és öngyilkossági rizikó szűrésének jelentősége, *Neuropsychopharmacologia Hungarica XVI. évf. 2. szám*.
21. Schwartz, H. A., Eichstaedt, J., Kern, M., Park, G., Sap, M., Stillwell, D., ... & Ungar, L. (2014, June). Towards assessing changes in degree of depression through facebook. In *Proceedings of the workshop on computational linguistics and clinical psychology: from linguistic signal to clinical reality* (pp. 118-125).
22. Szelei, A., & Döme, P. (2020): Daganatos megbetegedések és a depresszió: rövid irodalmi áttekintés, *Orvosi Hetilap*, 161(22), 908-916.
23. Titov, N., Dear, B., McMillan, D., Anderson, T., Zou, J. & Sunderland, M. (2011). Psychometric comparison of the PHQ 9 and BDI II for measuring response during treatment of depression, *Cogn BehavTherapy*, 40, 126-136.
24. Tringer, L. (2003). A depresszió: kor-kór?, *Mindentudás Egyeteme*.
25. Torzsa, P., Rihmer, Z., Gonda, X., Sebestyén, B., Szokontor, N., & Kalabay, L. (2008). A depresszió prevalenciája az alapellátásban Magyarországon, *Neuropsychopharmacologia Hungarica X/5*, 265-270.

26. World Health Organization (2020): <https://www.who.int/news-room/factsheets/detail/depression> (Letöltés dátuma: 2020.09.14.)
27. Yu-Ning Peng, Mei-Li Huang, & Chia-Hung Kao (2019). Prevalence of Depression and Anxiety in Colorectal Cancer Patients, International Journal of Environmental Research and Public Health. doi: [10.3390/ijerph16030411](https://doi.org/10.3390/ijerph16030411)
28. Zoltán Kmetty, & Renáta Németh (2020). Which is your favorite music genre? A validity comparison of Facebook data and survey data. doi: <https://arxiv.org/abs/2002.00501>
29. Zoltán Kmetty, & Károly Bozsonyi (2020): Identifying depression-related behavior on Facebook –an experimental study, Kézirat.

Melléklet:

```
Call:
glm(formula = depr_dummy ~ regg_ar + ferfi + harminc_alatt, family = "binomial",
    data = activ)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.7851 -0.7816 -0.5563 -0.3651  2.3733

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.7650     0.5254  -3.359 0.000781 ***
regg_ar      -0.7015     2.0722  -0.339 0.734949
ferfi        -0.8789     0.6661  -1.320 0.186979
harminc_alatt  0.7459     0.5501   1.356 0.175066
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 136.43  on 137  degrees of freedom
Residual deviance: 130.74  on 134  degrees of freedom
(12 observations deleted due to missingness)
AIC: 138.74

Number of Fisher Scoring iterations: 5
```

25. Ábra A depresszió diagnózisa és a reggeli aktivitás közti kapcsolat vizsgálata logisztikus regresszióval

```
Call:
glm(formula = depr_dummy ~ déle_ar + ferfi + harminc_alatt, family = "binomial",
    data = activ)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.8046 -0.7699 -0.5505 -0.3626  2.3690

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.7362     0.5829  -2.979 0.00289 **
déle_ar      -0.2595     0.9806  -0.265 0.79133
ferfi        -0.8517     0.6699  -1.271 0.20356
harminc_alatt  0.7745     0.5494   1.410 0.15860
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 136.43  on 137  degrees of freedom
Residual deviance: 130.79  on 134  degrees of freedom
(12 observations deleted due to missingness)
AIC: 138.79

Number of Fisher Scoring iterations: 5
```

26. Ábra A depresszió diagnózisa és a délelőtti aktivitás közti kapcsolat vizsgálata logisztikus regresszióval

```
Call:
glm(formula = depr_dummy ~ délu_ar + ferfi + harminc_alatt, family = "binomial",
    data = activ)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.8905 -0.7491 -0.5602 -0.3668  2.3484

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.4605     0.6784  -2.153 0.0313 *
délu_ar      -0.8912     1.1619  -0.767 0.4431
ferfi        -0.8750     0.6690  -1.308 0.1909
harminc_alatt  0.7402     0.5510   1.343 0.1791
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 136.43  on 137  degrees of freedom
Residual deviance: 130.26  on 134  degrees of freedom
(12 observations deleted due to missingness)
AIC: 138.26

Number of Fisher Scoring iterations: 5
```

27. Ábra A depresszió diagnózisa és a délutáni aktivitás közti kapcsolat vizsgálata logisztikus regresszióval

```

Call:
lm(formula = kogn_aff ~ regg_ar + ferfi + harminc_alatt, data = activ)

Residuals:
    Min       1Q   Median       3Q      Max
-1.5671 -0.6007 -0.2071  0.4644  2.5251

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.25090   0.15735  -1.595  0.1132
regg_ar      -0.93324   0.65674  -1.421  0.1576
ferfi        -0.03669   0.18291  -0.201  0.8413
harminc_alatt 0.41111   0.16646   2.470  0.0148 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8774 on 134 degrees of freedom
(12 observations deleted due to missingness)
Multiple R-squared:  0.06904, Adjusted R-squared:  0.0482
F-statistic: 3.313 on 3 and 134 DF, p-value: 0.02207

```

28. Ábra A kognitív-affektív skála és a reggeli aktivitás közti kapcsolat vizsgálata lineáris regresszióval

```

Call:
lm(formula = kogn_aff ~ déle_ar + ferfi + harminc_alatt, data = activ)

Residuals:
    Min       1Q   Median       3Q      Max
-1.5062 -0.5920 -0.2443  0.4920  2.5385

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.39116   0.18220  -2.147  0.0336 *
déle_ar      0.23459   0.33445   0.701  0.4843
ferfi        -0.05313   0.18647  -0.285  0.7761
harminc_alatt 0.42873   0.16681   2.570  0.0113 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8824 on 134 degrees of freedom
(12 observations deleted due to missingness)
Multiple R-squared:  0.05847, Adjusted R-squared:  0.03739
F-statistic: 2.774 on 3 and 134 DF, p-value: 0.04392

```

29. Ábra A kognitív-affektív skála és a délelőtti aktivitás közti kapcsolat vizsgálata lineáris regresszióval

```

Call:
lm(formula = kogn_aff ~ délu_ar + ferfi + harminc_alatt, data = activ)

Residuals:
    Min       1Q   Median       3Q      Max
-1.5315 -0.5817 -0.2282  0.4786  2.5376

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.29897   0.21568  -1.386  0.168
délu_ar      -0.05071   0.38436  -0.132  0.895
ferfi        -0.03213   0.18429  -0.174  0.862
harminc_alatt 0.43592   0.16684   2.613  0.010 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.884 on 134 degrees of freedom
(12 observations deleted due to missingness)
Multiple R-squared:  0.05514, Adjusted R-squared:  0.03398
F-statistic: 2.606 on 3 and 134 DF, p-value: 0.05436

```

30. Ábra A kognitív-affektív skála és a délutáni aktivitás közti kapcsolat vizsgálata lineáris regresszióval

```

Call:
lm(formula = kogn_aff ~ est_ar + ferfi + harminc_alatt, data = activ)

Residuals:
    Min       1Q   Median       3Q      Max
-1.5593 -0.5885 -0.2249  0.4393  2.5306

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.36105   0.18059  -1.999  0.04760 *
est_ar        0.20516   0.48796   0.420  0.67484
ferfi        -0.01484   0.18835  -0.079  0.93734
harminc_alatt 0.43685   0.16661   2.622  0.00975 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8834 on 134 degrees of freedom
(12 observations deleted due to missingness)
Multiple R-squared:  0.05626, Adjusted R-squared:  0.03513
F-statistic: 2.663 on 3 and 134 DF, p-value: 0.05061

```

31. Ábra A kognitív-affektív skála és az esti aktivitás közti kapcsolat vizsgálata lineáris regresszióval

```

Call:
lm(formula = szomat ~ regg_ar + ferfi + harminc_alatt, data = activ)

Residuals:
    Min       1Q   Median       3Q      Max
-1.9090 -0.6082 -0.0466  0.6115  2.1355

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.1130   0.1526  -0.740  0.4605
regg_ar      -0.1876   0.6371  -0.294  0.7689
ferfi       -0.3657   0.1774  -2.061  0.0412 *
harminc_alatt 0.2891   0.1615   1.790  0.0757 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8512 on 134 degrees of freedom
(12 observations deleted due to missingness)
Multiple R-squared:  0.07414, Adjusted R-squared:  0.05341
F-statistic: 3.577 on 3 and 134 DF, p-value: 0.01574

```

32. Ábra A szomatikus skála és a reggeli aktivitás közti kapcsolat vizsgálata lineáris regresszióval

```

Call:
lm(formula = szomat ~ dele_ar + ferfi + harminc_alatt, data = activ)

Residuals:
    Min       1Q   Median       3Q      Max
-1.89059 -0.62855 -0.04874  0.64511  2.11778

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.1681   0.1757  -0.957  0.3404
dele_ar       0.1350   0.3225   0.419  0.6761
ferfi       -0.3771   0.1798  -2.097  0.0378 *
harminc_alatt 0.2896   0.1608   1.800  0.0741 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8509 on 134 degrees of freedom
(12 observations deleted due to missingness)
Multiple R-squared:  0.07475, Adjusted R-squared:  0.05403
F-statistic: 3.608 on 3 and 134 DF, p-value: 0.01511

```

33. Ábra A szomatikus skála és a délelőtti aktivitás közti kapcsolat vizsgálata lineáris regresszióval

```

Call:
lm(formula = szomat ~ délu_ar + ferfi + harminc_alatt, data = activ)

Residuals:
    Min       1Q   Median       3Q      Max
-1.88822 -0.59464 -0.06079  0.67166  2.10967

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.000839   0.207178  -0.004  0.9968
délu_ar     -0.314062   0.369215  -0.851  0.3965
ferfi       -0.368458   0.177030  -2.081  0.0393 *
harminc_alatt 0.288726   0.160267   1.802  0.0739 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8491 on 134 degrees of freedom
(12 observations deleted due to missingness)
Multiple R-squared:  0.07851, Adjusted R-squared:  0.05788
F-statistic: 3.806 on 3 and 134 DF, p-value: 0.01174

```

34. Ábra A szomatikus skála és a délutáni aktivitás közti kapcsolatot vizsgálata lineáris regresszióval

```

Call:
lm(formula = szomat ~ est_ar + ferfi + harminc_alatt, data = activ)

Residuals:
    Min       1Q   Median       3Q      Max
-1.91455 -0.59690 -0.03899  0.61919  2.14297

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.1936    0.1737  -1.114  0.2672
est_ar       0.3285    0.4694   0.700  0.4853
ferfi       -0.3380    0.1812  -1.865  0.0643 .
harminc_alatt 0.2943    0.1603   1.836  0.0686 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8499 on 134 degrees of freedom
(12 observations deleted due to missingness)
Multiple R-squared:  0.07691, Adjusted R-squared:  0.05624
F-statistic: 3.722 on 3 and 134 DF, p-value: 0.01308

```

35. Ábra A szomatikus skála és az esti aktivitás közti kapcsolatot vizsgálata lineáris regresszióval

```

Call:
lm(formula = szomat ~ éjsz_ar + ferfi + harminc_alatt, data = activ)

Residuals:
    Min       1Q   Median       3Q      Max
-1.90238 -0.61359 -0.04466  0.61565  2.12135

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.1251    0.1485  -0.843  0.4009
éjsz_ar     -0.1010    1.8870  -0.054  0.9574
ferfi       -0.3640    0.1779  -2.046  0.0427 *
harminc_alatt 0.2953    0.1618   1.825  0.0703 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8514 on 134 degrees of freedom
(12 observations deleted due to missingness)
Multiple R-squared:  0.07356, Adjusted R-squared:  0.05282
F-statistic: 3.546 on 3 and 134 DF, p-value: 0.01636

```

36. Ábra A szomatikus skála és az éjszakai aktivitás közti kapcsolatot vizsgálata lineáris regresszióval

```

Call:
lm(formula = kogn_aff ~ munka + ferfi + harminc_alatt, data = ad_interest_full)

Residuals:
    Min       1Q   Median       3Q      Max
-1.6323 -0.6065 -0.2403  0.5290  2.4866

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.255095   0.184174  -1.385   0.1681
munka        -0.006008   0.008872  -0.677   0.4994
ferfi         0.039049   0.173237   0.225   0.8220
harminc_alatt 0.476825   0.159057   2.998   0.0032 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8851 on 146 degrees of freedom
Multiple R-squared:  0.06546, Adjusted R-squared:  0.04626
F-statistic: 3.409 on 3 and 146 DF, p-value: 0.01928

```

37. Ábra A kognitív-affektív skála és a munka témájú érdeklődési kategória közti kapcsolat vizsgálata lineáris regresszióval

```

Call:
glm(formula = depr_dummy ~ tarsas_kapcs + ferfi + harminc_alatt,
    family = "binomial", data = ad_interest_full)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.9531 -0.7649 -0.5385 -0.3679  2.3628

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.49172    0.63495  -2.349   0.0188 *
tarsas_kapcs -0.04572    0.03908  -1.170   0.2420
ferfi        -0.87070    0.61429  -1.417   0.1564
harminc_alatt 0.93809    0.53820   1.743   0.0813 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 150.12  on 149  degrees of freedom
Residual deviance: 142.10  on 146  degrees of freedom
AIC: 150.1

Number of Fisher Scoring iterations: 5

```

38. Ábra A depresszió diagnózisa és a társas kapcsolatok témájú érdeklődési kategória közti kapcsolat vizsgálata logisztikus regresszióval

```

Call:
lm(formula = kogn_aff ~ tarsas_kapcs + ferfi + harminc_alatt,
    data = ad_interest_full)

Residuals:
    Min       1Q   Median       3Q      Max
-1.5634 -0.6058 -0.2309  0.5992  2.4986

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.15453    0.20200  -0.765   0.44551
tarsas_kapcs -0.01635    0.01307  -1.251   0.21284
ferfi        -0.03586    0.18224  -0.197   0.84429
harminc_alatt 0.47368    0.15814   2.995   0.00322 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8818 on 146 degrees of freedom
Multiple R-squared:  0.07247, Adjusted R-squared:  0.05341
F-statistic: 3.803 on 3 and 146 DF, p-value: 0.01161

```

39. Ábra A kognitív-affektív skála és a társas kapcsolatok témájú érdeklődési kategória közti kapcsolat vizsgálata lineáris regresszióval


```

Call:
lm(formula = szomat ~ tarsas_kapcs + ferfi + harminc_alatt, data = ad_interest_full)

Residuals:
    Min       1Q   Median       3Q      Max
-1.86228 -0.62873 -0.03064  0.66114  2.03358

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.05223    0.19134   0.273  0.7853
tarsas_kapcs -0.01856    0.01238  -1.500  0.1359
ferfi        -0.41886    0.17262  -2.427  0.0165 *
harminc_alatt 0.33614    0.14979   2.244  0.0263 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8352 on 146 degrees of freedom
Multiple R-squared:  0.0944, Adjusted R-squared:  0.07579
F-statistic: 5.073 on 3 and 146 DF, p-value: 0.00227

```

40. Ábra A szomatikus skála és a társas kapcsolatok témájú érdeklődési kategória közti kapcsolat vizsgálata lineáris regresszióval

```

Call:
glm(formula = depr_dummy ~ vallas + ferfi + harminc_alatt, family = "binomial",
    data = ad_interest_full)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.8085 -0.7886 -0.5382 -0.3692  2.3090

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.89998    0.54682  -3.475 0.000512 ***
vallas      -0.05816    0.19249  -0.302 0.762517
ferfi       -0.69371    0.59226  -1.171 0.241483
harminc_alatt 0.94947    0.53854   1.763 0.077895 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 150.12 on 149 degrees of freedom
Residual deviance: 143.42 on 146 degrees of freedom
AIC: 151.42

Number of Fisher scoring iterations: 4

```

41. Ábra A depresszió diagnózisa és a vallás témájú érdeklődési kategória közti kapcsolat vizsgálata logisztikus regresszióval

```

Call:
lm(formula = kogn_aff ~ vallas + ferfi + harminc_alatt, data = ad_interest_full)

Residuals:
    Min       1Q   Median       3Q      Max
-1.6448 -0.5846 -0.2386  0.5404  2.5037

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.31033    0.16217  -1.914  0.0576 .
vallas      -0.01907    0.06222  -0.307  0.7596
ferfi        0.03162    0.17447   0.181  0.8564
harminc_alatt 0.47993    0.16009   2.998  0.0032 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8862 on 146 degrees of freedom
Multiple R-squared:  0.06313, Adjusted R-squared:  0.04388
F-statistic: 3.279 on 3 and 146 DF, p-value: 0.02278

```

42. Ábra A kognitív-affektív skála és a vallás témájú érdeklődési kategória közti kapcsolat vizsgálata lineáris regresszióval

```

Call:
glm(formula = depr_dummy ~ spiritualitas + ferfi + harminc_alatt,
     family = "binomial", data = ad_interest_full)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.8064 -0.7895 -0.5461 -0.3693  2.3321

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.90084    0.56114   -3.387 0.000705 ***
spiritualitas -0.04952    0.18958   -0.261 0.793918
ferfi        -0.70070    0.59601   -1.176 0.239735
harminc_alatt 0.94435    0.54265    1.740 0.081816 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 150.12  on 149  degrees of freedom
Residual deviance: 143.45  on 146  degrees of freedom
AIC: 151.45

Number of Fisher Scoring iterations: 4

```

43. Ábra A depresszió diagnózisa és a spiritualitás témájú érdeklődési kategória közti kapcsolat vizsgálata logisztikus regresszióval

```

Call:
lm(formula = kogn_aff ~ spiritualitas + ferfi + harminc_alatt,
    data = ad_interest_full)

Residuals:
    Min       1Q   Median       3Q      Max
-1.6441 -0.6076 -0.2242  0.5641  2.5039

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.29995    0.16884   -1.777 0.07773 .
spiritualitas -0.02395    0.06357   -0.377 0.70691
ferfi         0.02626    0.17592    0.149 0.88155
harminc_alatt 0.47422    0.16194    2.928 0.00396 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8861 on 146 degrees of freedom
Multiple R-squared:  0.06344, Adjusted R-squared:  0.04419
F-statistic: 3.296 on 3 and 146 DF, p-value: 0.02228

```

44. Ábra A kognitív-affektív skála és a spiritualitás témájú érdeklődési kategória közti kapcsolat vizsgálata lineáris regresszióval

```

Call:
lm(formula = szomat ~ spiritualitas + ferfi + harminc_alatt,
    data = ad_interest_full)

Residuals:
    Min       1Q   Median       3Q      Max
-2.04413 -0.58679 -0.03609  0.64180  2.01635

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.22535    0.15995   -1.409 0.1610
spiritualitas 0.05281    0.06022    0.877 0.3819
ferfi        -0.31106    0.16666   -1.866 0.0640 .
harminc_alatt 0.37813    0.15342    2.465 0.0149 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8394 on 146 degrees of freedom
Multiple R-squared:  0.08527, Adjusted R-squared:  0.06648
F-statistic: 4.537 on 3 and 146 DF, p-value: 0.004515

```

45. Ábra A szomatikus skála és a spiritualitás témájú érdeklődési kategória közti kapcsolat vizsgálata lineáris regresszióval

```

Call:
glm(formula = depr_dummy ~ egeszseg + ferfi + harminc_alatt,
     family = "binomial", data = ad_interest_full)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.9337 -0.7532 -0.5577 -0.3647  2.2563

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.44306   0.60427  -2.388  0.0169 *
egeszseg    -0.06379   0.04458  -1.431  0.1524
ferfi       -0.92261   0.61309  -1.505  0.1324
harminc_alatt 0.83849   0.54485   1.539  0.1238
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 150.12  on 149  degrees of freedom
Residual deviance: 141.27  on 146  degrees of freedom
AIC: 149.27

Number of Fisher Scoring iterations: 5

```

46. Ábra A depresszió diagnózisa és az egészség témájú érdeklődési kategória közti kapcsolat vizsgálata logisztikus regresszióval

```

Call:
lm(formula = szomat ~ egeszseg + ferfi + harminc_alatt, data = ad_interest_full)

Residuals:
    Min       1Q   Median       3Q      Max
-1.84625 -0.62170 -0.06107  0.67939  2.08657

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.01539   0.17926   0.086  0.9317
egeszseg    -0.01828   0.01294  -1.413  0.1598
ferfi       -0.41729   0.17349  -2.405  0.0174 *
harminc_alatt 0.30700   0.15278   2.009  0.0463 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.836 on 146 degrees of freedom
Multiple R-squared:  0.09286, Adjusted R-squared:  0.07422
F-statistic: 4.982 on 3 and 146 DF, p-value: 0.002551

```

47. Ábra A szomatikus skála és az egészség témájú érdeklődési kategória közti kapcsolat vizsgálata lineáris regresszióval