



TÁRSADALOMTUDOMÁNYI ADATFELDOLGOZÁS ÉS ELEMZÉS

**GEOGRÁFUS MESTERSZAK
(NAPPALI MUNKAREND)**

TANTÁRGYI KOMMUNIKÁCIÓS DOSSZIÉ

**MISKOLCI EGYETEM
MŰSZAKI FÖLDTUDOMÁNYI KAR
FÖLDRAJZ-GEOINFORMATIKA INTÉZET**

Miskolc, 2022

TARTALOMJEGYZÉK

1. Tantárgyleírás
2. Részletes tematika
3. Gyakorlati feladatok, zárthelyi dolgozat
4. Kompetencia táblázat

1. Tantárgyleírás

Tantárgy neve: Társadalomtudományi adatfeldolgozás és elemzés	Tantárgy kódja: MFKST730004
Tárgyfelelős: Siskáné Dr. Szilasi Beáta, egyetemi docens	Tárgyfelelős tanszék/intézet: Földrajz-Geoinformatika Intézet
Javasolt félév: 3.	Tantárgyelem: kötelező
Óraszám/hét (ea+gyak): 0+2	Előfeltételek: MFKHT71002, MFKFT720002
Kreditpont: 3	Számonkérés módja (a/gy/v): aláírás és gyakorlati jegy
Tagozat: nappali	
Tantárgy feladata és célja: A tárgy célja: a statisztikai alapfogalmak, a feldolgozandó adatok előkészítésével kapcsolatos eljárások áttekintése, az adatelemzés alapszintű elsajátítása statisztikai módszerek alkalmazásával.	
Fejlesztendő kompetenciák: tudás: Ismeri a földrajz vizsgálható folyamatait, rendszereit, tudományos problémáit, erről széles körű szakirodalmi tájékozottsággal rendelkezik. Ismeri a természet- és társadalomföldrajz mélyebb összefüggéseit, az ezekre vonatkozó elméleteket. Ismeri a földrajzi szakterületének sajátos kutatási (ismeretszerzési és problémamegoldási) módszereit, absztrakciós technikáit, az elvi kérdések gyakorlati vonatkozásainak kidolgozási módjait. Ismeri azokat a bonyolultabb terepi, laboratóriumi és gyakorlati módszereket, anyagokat és eszközöket, amelyekkel a geográfia területén munkáját gyakorolni tudja. Ismeri a geográfus alap- és alkalmazott kutatások tervezési és értékelési módjait. Ismeri a rendezett adatbázisokban elvégezhető műveleteket és modelleket. képesség: Képes a földrajz szakterületének sajátos kutatási módszereit, absztrakciós technikáit a gyakorlatban is alkalmazni a földrajztudományon belüli szűkebb szakterületén alternatív megoldások kidolgozására. Képes magas színvonalon alkalmazni az elsajátított terepi, laboratóriumi és gyakorlati módszereket, és speciális eszközöket. Képes földrajzi kutatások tervezésére, szervezésére, lebonyolítására és kutatások menedzselésére az eredményeit alkalmazó és továbbfejlesztő munkahelyeken, kutató- fejlesztő intézetekben és a szakigazgatásban. Képes a gyűjtött adatokat adatbázis rendszerbe rendezni, azokban különféle műveleteket végezni, modelleket alkotni, vagy a beszerzett adatokat geoinformatikai alapon rendszerezni, megjeleníteni. attitűd: Törekszik arra, hogy a földrajzi problémákkal kapcsolatos feladatait beosztott kollégáival együttműködve, szakmai véleményük figyelembevételével végezze. Munkája során kezdeményező szerepet vállal, szakmájának eredményeit a közösség szolgálatába állítja. Fejlett szakmai identitással, hivatástudattal rendelkezik. Törekszik arra, hogy a geográfia területén tudását folyamatosan továbbfejlessze. autonómia és felelősség: Jelentős mértékű önállósággal és felelősséggel végzi átfogó és speciális földrajzi szakképzettséget igénylő elméleti és gyakorlati összefüggések megalkotását, modellezését. Különböző bonyolultságú és különböző mértékben kiszámítható kontextusokban a módszerek és technikák széles körét alkalmazza önállóan a gyakorlatban.	
Tantárgy tematikus leírása: 1. Az adatfeldolgozással kapcsolatos általános fogalmak, megismerési módok, a kutatás időbelisége, az adatfeldolgozás helye a kutatás folyamatában. Konceptualizálás, operacionalizálás, mérési szintek. 2. Statisztikai alapfogalmak, minta, mintavétel, viszonyszámok, statisztikai táblák típusai. 3. Adatelőkészítés lépései, hibaellenőrzés, adattömörítés. Bevezetés a statisztikai elemzésekbe, leíró statisztika, egy-, és többváltozós elemzések. Gyakoriság, relatív gyakoriság.	

4. Középértékek (számtani átlag, átlag típusok, medián, kvantilisek, módusz). A szóródás mutatói (terjedelem, szórás, tapasztalati szórás, variancia, relatív szórás). Az eloszlás fogalma, nevezetes eloszlások, alakmutató számok.
 5. Változók közötti kapcsolat mérése. Asszociációs kapcsolat minőségi (nominális) változók között (χ^2 , Cramer-, és Csaprov féle asszociációs együtthatók, a PRE eljárás, λ , alternatív változók).
 6. Asszociációs kapcsolat mérése ordinális változók között (γ , asszociációs mátrix). Vegyes kapcsolat mennyiségi és minőségi változók között (teljes/belső/külső eltérés négyzetösszeg).
 7. Kapcsolaterősség mennyiségi változók között (regressziós függvény, pontdiagram, lineáris (Pearson-féle) korrelációs együttható, determinációs együttható).
 8. Függő és független változók, asszociációs mérőszám és oksági kapcsolat két változóra, a harmadik változó bevonása.
 9. Indexek és skálák, hasonlóságok és különbségek. Indexszerkesztés, szempontok, adathiányok kezelése, érvényesség ellenőrzése, külső itemellenőrzés. Skálaszerkesztés, szempontok, értékelésérvényesség. Bogardus-féle társadalmi távolság skála.
 10. Thurstone-skála, Likert-skála, Szemantikus differenciálskála, Guttman-skálák.
- Tipológiák
11. Lazarsfeld-módszer. Parciális kapcsolatok, nulladrendű kapcsolatok. Replikáció, explanáció, interpretáció, specifikáció. A Lazarsfeld-módszer finomítása.
 12. Hipotézisvizsgálat. Alapfogalmak, próbafüggvény, elfogadási tartomány, kritikus tartomány (bal-, jobb-, kétoldali). Szignifikanciaszint.
 13. A hipotézisvizsgálat lépései. Néhány egyszerű próba (z-próba, t-próba).
 14. Összegzés, záródolgozat.

Félévközi számonkérés módja:

A gyakorlatokon való részvétel az aláírás megszerzésének feltétele, maximális hiányzás 3 alkalommal. A gyakorlati jegy a zárthelyi dolgozat érdemjegye.

Értékelés:

100 - 85%: jeles; 75 – 84%: jó; 63 – 74%: közepes; 50 – 62%: elégséges; 0 - 50%: elégtelen.

Kötelező irodalom:

Babbie, E. 1999: A társadalomtudományi kutatás módszertana, Balassi Kiadó, 704 p.
 Falus I. – Ollé J. 2000: Statisztikai módszerek pedagógusok számára. Okker Kiadói Kft. 372 p.
 Hunyadi L. – Mundruczó Gy. – Vita L. 1996: Statisztika, Aula K, Budapest, 883 p.
 Rogerson, P.A. 2014: Statistical Methods for Geography : A Student's Guide. SAGE Publications Ltd. 4th Edition. 424 p.

Ajánlott irodalom:

Sajtos L. – Mitev A. 2007: SPSS kutatási és adatelemzési kézikönyv, Alinea Kiadó, Budapest, 402 p.

IBM SPSS Statistics 24 Documentation. IBM Support.
<https://www-01.ibm.com/support/docview.wss?uid=swg27047033#en>

2. Részletes tematika

Dátum	Gyakorlat
09. 07.	Az adatfeldolgozással kapcsolatos általános fogalmak, konceptualizálás, operacionalizálás, mérési szintek. Statisztikai alapfogalmak, minta, mintavétel, viszonyszámok, statisztikai táblák típusai.
09. 14.	Adatelőkészítés lépései, hibaellenőrzés, adattömörítés. Bevezetés a statisztikai elemzésekbe, leíró statisztika, egy-, és többváltozós elemzések. Gyakoriság, relatív gyakoriság.
09. 21.	Középértékek (számtani átlag, átlag típusok, medián, kvantilisok, módusz). A szóródás mutatói (terjedelem, szórás, tapasztalati szórás, variancia, relatív szórás). Az eloszlás fogalma, nevezetes eloszlások, alakmutató számok.
09. 28.	Elmarad (sportnap)
10. 05.	Elmarad (tanulmányi kirándulás)
10. 12.	Változók közötti kapcsolat mérése. Asszociációs kapcsolat minőségi (nominális) változók között (, Cramer-, és Csuprov féle asszociációs együtthatók, a PRE eljárás, λ , alternatív változók).
10. 19.	Asszociációs kapcsolat mérése ordinális változók között (γ , asszociációs mátrix).
10. 26.	Vegyes kapcsolat mennyiségi és minőségi változók között (teljes/belső/külső eltérés négyzetösszeg).
11. 02.	Kapcsolaterősség mennyiségi változók között (regressziós függvény, pontdiagram, lineáris (Pearson-féle) korrelációs együttható, determinációs együttható). Független és függő változók, asszociációs mérőszám és oksági kapcsolat két változóra, a harmadik változó bevonása.
11. 09.	Indexek és skálák, hasonlóságok és különbségek. Indexszerkesztés, szempontok, adathiányok kezelése, érvényesség ellenőrzése, külső itemellenőrzés. Skálaszerkesztés, szempontok, értékelésérvényesség. Bogardus-féle társadalmi távolság skála.
11. 16.	Thurstone-skála, Likert-skála, Szemantikus differenciálskála, Guttman-skálák. Tipológiák
11. 23.	Lazarsfeld-módszer. Parciális kapcsolatok, nulladrendű kapcsolatok. Replikáció, explanáció, interpretáció, specifikáció. A Lazarsfeld-módszer finomítása.
11. 30.	Hipotézisvizsgálat. Alapfogalmak, próbafüggvény, elfogadási tartomány, kritikus tartomány (bal-, jobb-, kétoldali). Szignifikanciaszint.
12. 07.	Összegzés, záródolgozat.

3. Gyakorlati feladatok, zárthelyi dolgozat

Társadalomtudományi adatfeldolgozás és elemzés záródolgozat

1. Mit nevezünk operacionalizálásnak? [1pont]

Azoknak az eljárásoknak a leírása, amikkel egy fogalom (ismérv) mérését, megfigyelését elvégezhetjük.

2. Milyen mérési szinteket különböztetünk meg? [4]

a, nem metrikus:

- nominális
- ordinális

b, metrikus

- intervallum
- arányskála

3. Írja fel az x változó standardizáltját! [1]

$$x'_i = (x_i - \bar{x})/s \Rightarrow \bar{x}' = 0, s' = 1, \text{ dimenziómentes}$$

4. Mi a módusz, medián, kvartilis? [3]

Módusz: a leggyakoribb érték

Medián: a sorrendbe állított értéksor középső értéke.

Kvartilis: a sorrendbe állított értékek negyedelése. Egy negyedbe az értékek 25%-a kerül.

5. Írja fel a korrigált tapasztalati szórásnégyzet képletét az $\{x_i\}$ ($i=1,2,\dots,n$) mintára! [2]

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

6. Állapítsa meg van-e kapcsolat a tanulmányi eredmény és az apa legmagasabb iskolai végzettsége között? (excel) [3]

	Az apa iskolai végzettsége			összes en
	8 ált.	érettségi	felsőfokú	
Tanulmányi eredmény < 3.0	12	8	4	

3.0-3.9	10	12	10	
>4.0	3	5	11	
összesen				

A táblában elvégezve az összegzéseket:

Tanulmányi eredmény	Az apa iskolai végzettsége			összesen
	8 ált.	érettségi	felsőfokú	
< 3.0	12	8	4	24
3.0-3.9	10	12	10	32
>4.0	3	5	11	19
összesen	25	25	25	75

Számítható: \square^2 , vagy \dagger , vagy \blacksquare

pl.:

$$E_1 = 75 - 25 = 50$$

$$E_2 = (24 - 12) + (32 - 12) + (19 - 11) = 12 + 20 + 8 = 40$$

$$\lambda = PRE = (E_1 - E_2) / E_1 = (50 - 40) / 50 = 0,2 = 20\%$$

A kapcsolat gyenge.

7. Írjon példát Guttman-skála vizsgálatra! (excel) [5]

pl.: Egy adott kérdésre adható A, B, C válaszokkal való egyetértés vizsgálatával a válaszok skálázottságára a következő eredményeket kaptuk (+: egyetért, - nem ért egyet):

	A	B	C	esetszám	indexpont	skálapontszám	Összes skálahiba
Skála-típusú válaszok	+	+	+	612	3	3	0
	+	+	-	448	2	2	0
	+	-	-	92	1	1	0
	-	-	-	79	0	0	0
Kevert típusú válaszok	-	+	-	15	1	2	15=15*(2-1)
	+	-	+	5	2	3	5=5*(3-2)
	-	-	+	2	1	3	4=2*(3-1)
	-	+	+	5	2	3	5=5*(3-2)

$$\text{reprodukálhatósági együttható} = 1 - \frac{\text{összeshiba}}{\text{összesválasz}}, \text{ itt } = 1 - \frac{29}{1258 * 3} = 0,99 = 99\%,$$

ahol 1258 a válaszadók száma, 29 az összes hiba. A 99% nagyon jó skálázhatóságot jelent

8. Írjon egy példát a Lazarsfeld-modell alkalmazására! (excel) [5]

PL.:

Eredeti változók:

A: gólyák száma a településen [kevés,sok]

B: születési arány a településen [alacsony,magas]

$A \Rightarrow B$ népi vélekedés: a gólyák hozzák a gyereket, így a gólyák száma meghatározza a születések számát

C Megelőző típusú kontrollváltozó: településtípus [falu,város]

$C \Rightarrow A$

$\Rightarrow B$

A és B kapcsolata:

<i>B</i>	<i>A</i>	
	<i>kevés</i>	<i>sok</i>
<i>Alacsony</i>	90 %	10 %
<i>magas</i>	10 %	90 %

A táblázatban leírt kapcsolat valóban azt mutatja, hogy sok gólya esetén magas a születési arányszám.

Kíváncsiak vagyunk arra, hogy a településtípus hatással van-e erre a kapcsolatra.

<i>B</i>	<i>C</i>			
	<i>falu</i>		<i>város</i>	
	<i>kevés</i>	<i>sok</i>	<i>kevés</i>	<i>sok</i>
<i>Alacsony</i>	0	0	100	100
			%	%
<i>Magas</i>	100	100	0	0
	%	%		
	100	100	100	100
	%	%	%	%
	<i>1. parciális</i>		<i>2. parciális</i>	

Egyik parciálisban sincs kapcsolat A és B között (A értéke szerinti változás nem változtatja B eloszlását). Az 1. parciális szerint függetlenül a gólyák számától a születési arány a falvakban magas, a 2. parciális szerint a városokban alacsony. Így tehát kijelenthetjük, hogy az A,B kapcsolatot teljesen a kontrollváltozó, a településtípus okozza.