

IFI6081.DT	ANDMEANALÜÜS 1		
Maht: 3EAP	Kontakttundide maht: 34 tundi	Õppesemester: K	Eksam
Eesmärk:	▪ Luua võimalused andmete kirjeldamise ja seoste statistilist usaldusväärsust peegeldava analüüsi läbiviimiseks teoreetiliste teadmiste ja praktiliste oskuste omandamise kaudu. ▪ Tutvustada tuntud statistikapaketi SPSS võimalusi kirjeldava ja üldistava statistika meetodite kasutamisel. ▪ Aidata kaasa teadmiste ja oskuste praktilise rakendamiskogemuse kujunemisele, mis võimaldab teha iseseisvalt otsustusi sobiva(te) analüüsimeetodi(te) valikuks ning analüüsi tulemusi korrektselt tõlgendada.		
Aine lühikirjeldus:	Teemad: ▪ Andmete/tunnuste tüübid. ▪ Esmane andmete analüüs. Andmete esitamine tekstina. Sagedustabelid. Andmete graafiline kirjeldamine. Kirjeldavad arvnäitajad. Normaaljaotuse idee. ▪ Üldkogum ja valim. Üldkogumi arvnäitajate hindamine - vahemikhinnang. Statistiline üldistamine. Parameetrilised ja mitteparameetrilised analüüsimeetodid. ▪ Mitme valikuvõimalusega küsimuste analüüs. ▪ Korrelatsioonanalüüs. Seosekordajate statistiline olulisus. Kursuse korraldus: Kursuse sooritamiseks vajalik töö maht on 78 (3x26) tundi. Sellest kontakttundidena toimuvad seminarid (28 tundi) ja eksam (6 tundi). Eksam koosneb kahest osast: kirjalikust avatud küsimustega testist ning iseseisvast tööst.		
Õpiväljundid:	▪ Eristab andmete/tunnuste tüüpe ning valib vastavalt andmete tüübile ning andmete kohta esitatud küsimuse sisule sobivad analüüsi meetodid (käsitletud meetodite piires). ▪ Omab kogemust andmetest lähtuvate ning statistilist kirjeldamist/üldistamist eeldavate küsimuste püstitamiseks. ▪ Mõistab käsitletud esmase analüüsi meetodite olemust, teab nende rakendamise tingimusi ning oskab analüüsi tulemusi korrektselt tõlgendada. ▪ Hindab üldkogumi arvnäitajaid, kasutades vastavalt lähteülesandele kas parameetrilisi või mitteparameetrilisi meetodeid. ▪ Selgitab näidete põhiselt parameetriliste ja mitteparameetriliste meetodite erinevusi. ▪ Oskab juhendmaterjali abiga kasutada vastavat tarkvara andmetöötluse ja üldistava statistika meetoditel põhineva analüüsi läbiviimiseks.		
Hindamis-meetodid:	Eksam (hindeline) Eksamihinne kujuneb kirjaliku avatud küsimustega testi (50%) ja iseseisvate tööde tulemuste (50%) põhjal. Positiivse hinde saamiseks on vajalik saada nii testi kui ka iseseisva töö tulemuseks vähemalt 51%.		
Õppejõud:	lekt Kairi Osula		
Inglisekeelne nimetus:	Data Analysis I		
Eeldusaine:	Arvutikasutuse elementaaroskused IFI6001mahus		
Kohustuslik kirjandus:	Niglas, K. Videoloengud andmeanalüüsist (saadaval informaatika instituudis) Osula, K. (2015) Kursuse materjalid (http://www.tlu.ee/~kairio/6081) Niglas, K. Statistika loengumaterjale (http://www.cs.tlu.ee/~katrin/wp/) Niglas, K. (2007) Andmeanalüüs statistikapaketi SPSS 14.00 abil. Põhikursus Tallinn, TLÜ.		
Asendus-kirjandus:	Hiob, K. (1995) Matemaatiline statistika. Algkursus koolidele, Tallinn Parring, A.-M., Vähi, M., Käärrik, E. (1997) Statistilise andmetöötluse algõpetus, Tartu Tooding, L.-M. (1999) Andmeanalüüs sotsiaalteadustes, Tartu Tooding, L.M. (2007) Andmete analüüs ja tõlgendamine sotsiaalteadustes, Tartu		
Õppetöös	Seminarides osalemine on kohustuslik.		

osalemise ja eksamile/arvestusele pääsemise nõuded:	Hindamiseks peab üliõpilane esitama kolmest osast koosneva iseseisva töö ja sooritama kirjaliku avatud küsimustega testi. Iseseisvate tööde esitamise tähtjad, nõuded ja sisu on toodud kursuseprogrammi ajakavas. Nii iseseisvat tööd kui testi võib esitada/sooritada kuni 2016 aasta sügissemestri vahenädalal väljakuulutatud järeleksamini. Hiljem tuleb aine uuesti kuulata.
Iseseisva töö nõuded:	Igal üliõpilasel tuleb kursuse käigus koostada kolmest osast koosnev iseseisev töö. Esimene ja teine töö (rühmatööd) koosnevad etteantud teemadel uuringu läbiviimisest, materjali läbitöötamisest ja kokkuvõtte tegemisest. Kolmas töö (individuaalne töö) koosneb õppejõu poolt ette antud praktilistest andmeanalüüsi ülesannetest kogu läbitud materjali ulatuses. Kasutatavad andmestikud võivad olla kas aine käigus kogutud, õppejõu poolt ette antud või üliõpilaste poolt mingi teise aine raames kogutud (nende kasutamine tuleb õppejõuga eelnevalt kooskõlastada). Kolmanda iseseisva töö esitamisel tuleb üliõpilasel oma tööd kaitsta – vastata tööga seonduvatele küsimustele ning põhjendada tehtud valikuid.
Eksami hindamis-kriteeriumid	Eksamihinne kujuneb kirjaliku avatud küsimustega testi (50%) ja iseseisvate tööde tulemuste (50%) koondtulemusena järgmiselt: „A” - suurepärane 91-100% „B” - väga hea 81- 90% „C” - hea 71- 80% „D” - rahuldav 61-70% „E” - kasin 51- 60% „F” - puudulik 0 – 50% Positiivse eksamihinde saamiseks peavad olema positiivsele tulemusele (vähemalt 51%) esitatud mõlemad osad (kirjalik test, is.töö), st. et ei piisa vaid testi või koduste ülesannete esitamisest. Põhieksami raames toimuval kirjalikul testil võib üliõpilane kasutada paberkandjal materjale, järeleksamil ei ole materjalide kasutamine lubatud. Kirjalik (avatud küsimustega) test moodustab 50% eksamihindest ning seda hinnatakse järgmiste kriteeriumide alusel: Teema: Statistilised andmed. Tunnusetüübid. (41-50%) Selgitab teemakohaseid põhimõisteid. Kasutab neid korrektselt praktiliste ülesannete lahendamisel tuues näiteid. Väga head taset ületav teadmiste ja oskuste vaba ning loov kasutamine. (1-10%) Tunnuse tüüpide määramisel ning vastavalt tunnuse tüübile sobiva andmeanalüüsi meetodi valimisel esineb eksimusi. Teema: Andmete kirjeldamise põhilised viisid (tekst, tabel, diagramm). (41-50%) Selgitab teemakohaseid põhimõisteid. Kasutab neid korrektselt praktiliste ülesannete lahendamisel tuues näiteid. Väga head taset ületav teadmiste ja oskuste vaba ning loov kasutamine. (1-10%) Andmetele vastava esitlusviisi valikul esineb eksimusi. Valitud viis ei ole mõningatel juhtudel sobivaim. Järelduse kirjutamisel esineb puudujääke. Teema: Kirjeldavad arvnäitajad. (41-50%) Selgitab teemakohaseid põhimõisteid. Kasutab neid korrektselt praktiliste ülesannete lahendamisel tuues näiteid. Väga head taset ületav teadmiste ja oskuste vaba ning loov kasutamine. (1-10%) Esitab kirjeldavate arvnäitajate definitsioonid. Mõistete kasutusoskus praktilises olukorras osaliselt puudulik. Teema: Üldkogum ja valim. Üldkogumi arvnäitajate hindamine - vahemikhinnang. Statistiline üldistamine. Parameetrilised ja mitteparameetrilised analüüsimeetodid.

	(41-50%) Selgitab teemakohaseid põhimõisteid. Kasutab neid korrektselt praktiliste ülesannete lahendamisel. Väga head taset ületav teadmiste ja oskuste vaba ning loov kasutamine. (1-10%) Esitab vahemikhinnangu valemid. Mõistete ja valemite kasutusoskus praktilises olukorras osaliselt puudulik. Teema: Korrelatsioonanalüüs. (41-50%) Selgitab teemakohaseid põhimõisteid. Kasutab neid korrektselt praktiliste ülesannete lahendamisel tuues näiteid. Väga head taset ületav teadmiste ja oskuste vaba ning loov kasutamine. (1-10%) Selgitab korrelatsioonanalüüsi sisu ning nimetab enamlevinud kordajad. Sobiva kordaja valmisel vastavalt andmete tüübile esineb eksimusi. Iseseisvad tööd moodustavad eksamihindest 50% ning neid hinnatakse järgmiselt: I töö (rühmatöö) 10% + II töö (rühmatöö) 15% + III töö (individuaalne töö) 25% I töö (rühmatöö) ülesandeid hinnatakse skaalal: Väga hea töö (8-10%) – uuringul on selge eesmärk ning koostatud küsimused vastavad eesmärgile. Küsimuste koostamisel on kasutatud täpselt nõutud tunnuse tüüpe ning valitud skaalad on korrektsed. Küsimused on üheseltmõistetavad ja arusaadavad. Hea töö (5-7%) – uuringu eesmärk on selge, koostatud küsimused vastavad üldjoontes püstitatud eesmärgile. Vastusevariantides esineb üksikuid vigu. Töö on arvestamata (0-4%) II töö (rühmatöö) ülesandeid hinnatakse skaalal: Väga hea töö (15-12%) – analüüs vastab teemale ning sellel on selge eesmärk. Meetodite kasutamine on kooskõlas analüüsi eesmärgi ja valitud tunnuste tüüpidega. Hea töö (11-8%) – analüüs on suures osas eesmärgipärane, meetodite kasutamisel esineb ebatäpsusi. Töö on arvestamata (0-7%) III töö (individuaalse töö) ülesandeid hinnatakse skaalal: Väga hea töö (25-22%) - koduste ülesannete lahendamisel on valitud probleemidele sobivaimad meetodid. Valikut on põhjendatud. Tulemid on korrektselt kujundatud ning töö on visuaalselt hästi loetav/haaratav. Järeldused on kirjutatud korrektselt viidates tulemile. Hea töö (21-18%) - koduste ülesannete lahendamisel on suures osas valitud sobivaimad andmete esitamise meetodid. Mõnede näidete puhul on valik põhjendamata. Töö on üldjoontes vormistatud korrektselt (tulemid kujundatud). Järelduste kirjutamisel esineb üksikuid puudujääke. Rahuldav töö (17-13%) - koduste ülesannete lahendamisel on puudu üksikud ülesanded. Meetodile sobivate probleemide püstitamisel esineb puudujääke. Tulemid on osaliselt kujundamata. Järeldused on suures osas õiged, kuid esineb üksikuid puudujääke. Töö on arvestamata ning tuleb esitada uuesti (0-12%).
--	--

Õppetöö sisu ja ajakava

<i>Kuupäev</i>	Teema, sisu lühikirjeldus
1. seminar 4.02.2016	Sissejuhatus ainesse. Ülevaade kursuse korraldusest. SPSS for Windows. Tunnuste defineerimine, andmete sisestamine.
Rühmatöö (1) 10p.	Ühe rühma moodustavad kuni 3 üliõpilast – miniuuringu läbiviimise I etapp: uurimisküsimuse püstitamine, küsimustiku loomine (10-12 küsimust). Tulemuste ettekandmine (vajalik kõikide rühmaliikmete kohalolu) toimub 3.seminaris (18.02). <i>Seminarist puudumise korral palun esitada (kogu) ettekande tekst kirjalikult.</i> <i>Tähtajast hiljem tehtud töö on personaalne töö ning maksimumpunktid 6p.</i>
2. seminar 11.02.2016	Põhilised andme- ja failiteisendused: sorteerimine, arvutamine, andmete jagamine gruppidesse, selekteerimine. Andmete ületoomine veebipõhistest andmekogumiskeskondadest SPSS-i. Andmete transport Excelist SPSS-i.
3. seminar 18.02.2016	1. rühmatöö tulemuste ettekandmine. Arutelu. Andmestiku korrastamine (analüüsiks ettevalmistus): puuduvate väärtuste analüüs ja asendamise võimalused, sisestusvigade otsimine, skaalade pööramine. Andmete tüübid.
4. seminar 25.02.2016	Andmestiku korrastamine. Ühemõõtmeline analüüs: andmete esitamine tekstina, sagedustabelid. Tulemuste redigeerimine ja tõlgendamine.
5. seminar 3.03.2016	Kirjeldavate arvnäitajate arvutamine, tulemuste tõlgendamine.
6. seminar 10.03.2016	Täiendavad võimalused kirjeldavate arvnäitajate leidmiseks. Andmete graafiline kirjeldamine (histogramm).
7. seminar 17.03.2016	Andmete graafiline kirjeldamine – tulpdiaagramm. Sektordiaagramm. Andmete kirjeldamise (arvnäitajad, tabelid, diagrammid) kordamine.
Rühmatöö (2) 15p.	Ühe rühma moodustavad kuni 3 üliõpilast – miniuuringu läbiviimise II etapp: tulemuste esmane analüüs. Andmestiku korrastamine ja lihtsamate õpitud statistiliste meetodite rakendamine. Tulemuste ettekandmine (vajalik kõikide rühmaliikmete kohalolu) toimub 8.seminaris (31.03). <i>Seminarist puudumise korral palun esitada (kogu) ettekande tekst kirjalikult.</i> <i>Tähtajast hiljem tehtud töö on personaalne töö ning maksimumpunktid 10p.</i>
24.03.2016	Vahenädal, tundi ei toimu. Iseseisva töö (2) tegemine.
8. seminar 31.03.2016	2. rühmatöö tulemuste ettekandmine. Arutelu.
9. seminar 7.04.2016	Parameetrilised ja mitteparameetrilised analüüsimeetodid. Normaaljaotuse idee, selle kasutamine statistikas. Üldkogum ja valim.
10. seminar 14.04.2016	Vahemikhinnang. Üldkogumi keskväärtuse ning proportsioonide hindamine.
11. seminar 21.04.2016	Vahemikhinnang. Vahemikhinnangu tulemuste graafiline esitamine. Korrelatsioonanalüüs.
12. seminar 28.04.2016	Korrelatsioonanalüüs. Seosekordajate statistiline olulisus. Mitme valikuvõimalusega küsimuste kirjeldamine.
13. seminar	Kordamine. Is.töö nr. 3. ülesannete tutvustamine.

5.05.2016	
Is.töö (3) 25p.	Personaalne töö – kogutud andmestike põhjal õpitud meetodite rakendamine. Tähtaeg: vastavalt tunnis kokkulepitud graafikule.
14. seminar 12.05.2016 10.15-13.45	Konsultatsioon. Is.töö nr. 3 kaitsmine.
Eksam 1 12.05.2016 14.15-15.45	Kirjalik test 14.15-15.45
Eksam 2 20.05.2016 10.15-13.45	Kirjalik test: 10.15-11.45 Is.töö nr. 3 kaitsmine 12.15-13.45