

Kursuseprogramm – IFI7044

IFI7044	ANDMEANALÜÜS: ÜLDISTAV STATISTIKA		
4 EAP	Kontakt tundide maht: 22	Õppesemester: S	Eksam
Eesmärk:	Luu võimalused erinevuste ja seoste statistilist usaldusväärsust peegeldava analüüsi läbiviimiseks kasutatavate teoreetiliste teadmiste ja praktiliste oskuste omandamiseks. Tutvustada tuntud statistikapaketi SPSS võimalusi üldistava statistika põhimeetodite kasutamisel. Aidata kaasa teadmiste ja oskuste ning praktilise rakendamiskogemuse kujunemisele, mis võimaldab õppijal teha iseseisvalt otsustusi sobiva(te) analüüsimeetodi(te) valikuks ning analüüsi tulemusi korrektselt tõlgendada.		
Aine lühikirjeldus: (sh iseseisva töö sisu kirjeldus vastavuses iseseisva töö mahule)	Teemad: Üldkogum ja valim. Normaaljaotuse idee. Parameetrilised ja mitteparameetrilised analüüsimeetodid. Statistiline üldistamine: üldkogumi arvnäitajate hindamine; statistilised olulisustestid: t-test, χ^2-test, ANOVA, Kruskal-Wallise test; seosekordajate statistiline olulisus. Statistilise olulisuse tõlgendamine. Sobiva analüüsi- ja esitlusmeetodi valik. Korraldus: Kursuse sooritamiseks vajalik töö maht on 104 (4x26) tundi. Sellest kontakt tundidenä toimuvad seminarid (20 tundi) ja eksam (2 tundi). Üliõpilased on kohustatud enne seminare (v.a. esimene seminar) vaatama programmikohased lõigud videoloengutest või tutvuma vastava teemaga kirjalike materjalide põhjal. Peale seminare on üliõpilased kohustatud kinnistama õpitu iseseisva õppimise teel ning lahendama 3 kodust ülesannet (vastavalt teemale, kas personaalsed või rühmatööd). Eksami läbimiseks tuleb sooritada kirjalik avatud küsimustega test. Testi hinnet mõjutab koduste ülesannete koondtulemus ühe punkti võrra. Vt. täpsemalt: „Eksami hindamiskriteeriumid“		
Õpiväljundid:	▪ Omab kogemust andmetest lähtuvate ning statistilist üldistamist eeldavate küsimuste püstitamiseks. ▪ Mõistab käsitletud üldistava statistika meetodite olemust, teab nende rakendamise tingimusi ning oskab analüüsi tulemusi korrektselt tõlgendada. ▪ Oskab eristada andmete/tunnuste tüüpe ning valida vastavalt andmete tüübile ning andmete kohta esitatud küsimuse sisule sobivad analüüsi meetodid (käsitletud meetodite piires). ▪ Oskab juhendmaterjali abiga kasutada vastavat tarkvara andmetöötluse ja üldistava statistika meetoditel põhineva analüüsi läbiviimiseks.		

Hindamismeetodid:	Eksam (hindeline). Hinne kujuneb iseseisva töö ja lahtiste küsimustega testi tulemuste põhjal. Positiivse hinde saamiseks on vajalik saada nii testi kui ka iseseisva töö tulemuseks vähemalt 51%. Hindamisele pääsevad kõik ainele registreeritud tudengid vaatamata sellele, kas ja kui palju on nad õppetöö kontakttundides osalenud. Vt täpsemalt kursuseprogrammi alalõik „Hindamiskriteeriumid“.
Õppejõud:	Kairi Osula
Ingliskeelne nimetus:	Data Analysis: Inferential Statistics
Eeldusaine:	Teadmised IFI7041 mahus
Kohustuslik kirjandus:	Niglas, K. Videoloengud andmeanalüüsist (saadaval informaatika instituudis) Osula, K. Kursuse materjalid (www.tlu.ee/~kairio/7044) Niglas, K. Statistika loengumaterjale (www.tlu.ee/~katrin/) Niglas, K. (2007) Andmeanalüüs statistikapaketi SPSS 14.00 abil. Põhikursus Tallinn, TLÜ.
Asenduskirjandus: (üliõpilase poolt läbi töötatava kirjanduse loetelu, mis katab ainekursuse loengulist osa)	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 osalemise ja eksamile/arvestusele pääsemise nõuded	Kontakttundides osalemine ei ole kohustuslik (eeldab seminari ajaks eelneva tunni materjali iseseisvat omandamist). NB! Hindamiseks peab õpilane esitama 3 kodust tööd ja sooritama testi. Nii koduseid töid kui testi võib esitada/sooritada ühe aasta jooksul peale kursuse lõppemist st 2012 aasta sügissemestril toimuva aine sooritused peavad olema tehtud/esitatud hiljemalt 2014 aasta kevadsemestri eelnädalaks välja kuulutatud korduseksami tähtajal/tähtajaks.
Iseseisva töö nõuded	Igal üliõpilasel tuleb kursuse käigus jooksvalt koostada 3 erinevat iseseisvat tööd. Iseseisevad tööd varieeruvad teemati ning on rakendusliku/praktilise iseloomuga. Viimane (3.) töö koosneb õppejõu poolt ette antud praktilistest andmeanalüüsi ülesannetest kogu läbitud materjali ulatuses. Kasutatavad andmestikud võivad olla kas õppejõu poolt ette antud või üliõpilaste poolt mingi teise aine raames kogutud (nende kasutamine tuleb õppejõuga eelnevalt kooskõlastada).

Eksami hindamiskriteeriumid	Eksami hinne kujuneb kirjaliku testi ja iseseisvatöö ülesannete koondtulemusena järgmiselt: 1. avatud küsimustega test: küsimused/ülesanded valitakse testi nii, et nad peegeldaks komplektina nelja esimest programmis kirjeldatud õpitulemust; iga küsimus/ülesanne annab teatud arvu punkte; punktid summeeritakse ning hinne kujuneb ülikoolis tavaks kujunenud süsteemi alusel: „A” (91-100%) – silmapaistev ja eriti laiapõhjaline õpiväljundite saavutamise tase, mida iseloomustab väga head taset ületav teadmiste ja oskuste vaba ning loov kasutamine; „B” (81-90%) – väga heal tasemel õpiväljundite saavutamine, mida iseloomustab teadmiste ja oskuste eesmärgipärane ja loov kasutamine. Spetsiifilisemate ja detailsemate teadmiste ja oskuste osas võivad ilmned mittesisulisel ja mittepõhimõttelised eksimused; „C” (71-80%) – heal tasemel õpiväljundite saavutamine, mida iseloomustab teadmiste ja oskuste eesmärgipärane kasutamine. Spetsiifilisemate ja detailsemate teadmiste ja oskuste osas avaldub ebakindlus ja ebatäpsus; „D” (61-70%) – piisaval tasemel õpiväljundite saavutamine, mida iseloomustab teadmiste ja oskuste kasutamine harjumuspärase olukorras, kuid erandlikes olukordades avalduvad puudujäägid ja ebakindlus; „E” (51-60%) – minimaalselt lubataval tasemel olulisemate õpiväljundite saavutamine, mida iseloomustab teadmiste ja oskuste kasutamine tüüpilistes olukordades piiratud viisidel, kuid erandlikes olukordades avalduvad märgatavad puudujäägid ning ebakindlus; „F” (alla 51%) – õppija on omandanud teadmised ja oskused miinimumtasemest madalamal tasemel. Hinne "F" on negatiivne tulemus ning test tuleb igal juhul uuesti sooritada. 2. iseseisvaid töid hinnatakse skaalal: „+“ – väga hea töö 85-100% (tõstab testi hinnet ühe hinde võrra) 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. „0“ – hea töö 67-84% (jätab testi hinde samaks) 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 51-66% (langetab testi hinnet ühe hinde võrra) 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. „F“ – töö arvestamata 0-50% (töö esitamata või esitatud mitterahuldavalt ning tuleb uuesti esitada) Eksami hinde saamiseks peavad olema positiivsele tulemusele tehtud mõlemad tööd (kirjalik test, iseseisvad tööd).
------------------------------------	---

Õppetöö sisu ja ajakava

Kuupäev	Teema, sisu lühikirjeldus
1.seminar 8.09.2012 14.00-18.00 T-302	Sissejuhatus kursusesse. Ülevaade kursuse korraldusest. Juhuslik valim, selle tähtsus ja moodustamise viisid. Statistiline viga. Valimikeskväärtuste jaotuse standardhälve e. standardviga. Üldkogumi keskväärtuse ja teiste arvnäitajate hindamine. Usaldusnivoo. Usaldusintervall. Usalduspiirid. Statistilised olulisustestid - statistiline hüpotees; selle kontrollimine. Olulisusnivoo ja olulisustõenäosus.
2. seminar 22.09.2012 14.00-18.00 T-302	Kordamine: statistilised olulisustestid - statistiline hüpotees; selle kontrollimine. Olulisusnivoo ja olulisustõenäosus. Kahe valimi keskväärtuste võrdlemine – t-test. Sõltuvad ja sõltumatud valimid. Sõltuvate valimite t-test. Sõltumatute valimite t-test. Ühe- ja kahepoolsed olulisustestid.
1. kodune töö Rühmatöö (ühes rühmas 2-4 õpilast). (25 punkti)	Töö sisu: * Kirjalik töö vahemikhinnangu arvutamisele ning saadud tulemuste tõlgendamisele. * Iseseisvalt olulisustestide kohta (lisa)materjali lugemine. Lühikokkuvõtte koostamine. * Etteantud t-testi tulemuste tõlgendamine. Töö esitamise tähtaeg: 20.10.2012 Tähtajast hiljem esitatud töö maksimumtulemus 10 punkti
3. seminar 20.10.2012 14.00-18.00 T-302	1.koduse töö tulemuste arutelu. Kordamine: sõltumatute valimite t-test. Dispersioonanalüüs. Post-Hoc testid.
4. seminar 3.11.2012 14.00-18.00 T-302	Mitteparameetrilised meetodid. Risttabelid ja χ^2 -test.
2. kodune töö Rühmatöö (ühes rühmas 2-4 õpilast). (25 punkti)	Töö sisu: * Iseseisvalt dispersioonanalüüsi ning MPA testide kohta (lisa)materjali lugemine. Rühmaarutelu. * Etteantud dispersioonanalüüsi ja hii-ruut testi tulemuste tõlgendamine. Töö esitamise tähtaeg: 1.12.2012 Tähtajast hiljem esitatud töö maksimumtulemus 10 punkti
5. seminar 1.12.2012 14.00-18.00 T-302	2.koduse töö tulemuste arutelu. Kordamine: Korrelatsioonanalüüs. Korrelatsioonikordaja statistiline olulisus. Üldine kordamine: kõikide õpitud meetodite rakendamine.
3. kodune töö Personaalne töö (50 punkti)	Töö sisu: Praktiliste andmeanalüüsi ülesannete lahendamine kasutades õpitud üldistava statistika meetodeid. Töö esitamise tähtaeg: 12.01.2013 Tähtajast hiljem esitatud töö maksimumtulemus 40 punkti
Kirjalik test	12.01.2013 kell 10.00-12.00

Õppeainet kureeriv üksus:	Informaatika instituut
Kursuseprogrammi koostaja	Kairi Osula
Allkiri:	
Kuupäev:	16.08.2012