

Kursuseprogramm

IFI7004	ANDMEANALÜÜS: ÜLDISTAV STATISTIKA		
4 EAP	Kontaktitudide maht: 22	Õppesemester: S	Eksam
Eesmärk:	Lua võimalused erinevuste ja seoste statistilist usaldusväärsust peegeldava analüüsi läbiviimiseks vana minevate 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 kontaktitudidena toimub sissejuhatav seminar (4 tundi), praktikumid (16 tundi) ja eksam (2 tundi). Üliõpilased on kohustatud enne praktikume vaatama programmikohased lõigud videoloengutest või tutvuma vastava teemaga kirjalike materjalide põhjal (20 tundi) ning peale praktikume kinnistama õpitu iseseisva õppimise teel (30 tundi). Igal üliõpilasel tuleb kursuse lõpus esitada mahukas arvestuslik iseseisev töö (32 tundi), mis koosneb õppejõu poolt ette antud praktilistest andmeanalüüsi ülesannetest kogu läbitud materjali ulatuses. Eksami läbimiseks tuleb sooritada ka kirjalik avatud küsimustega test.		
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. Kirjalikku testi hinnatakse õppekorralduse eeskirjas toodud alustel; iseseisev töö võib testi põhjal saadud hinnet ühe hinde võrra tõsta (suurepärane või väga hea töö), samaks jätta (hea või rahuldav töö) või ühe hinde võrra langetada (kasin töö). Vt ka alalõik „hindamiskriteeriumid“.
Õppejõud:	Prof Katrin Niglas
Inglisekeelne nimetus:	Data Analysis: Inferential Statistics
Eeldusaine:	teadmised IFI7041 mahus
Kohustuslik kirjandus:	Niglas, K. Videoloengud andmeanalüüsist (saadaval informaatika instituudis) Niglas, K. Statistika loengumaterjale (http://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 on väga soovitatav, kuid lubatud on ka täiesti iseseisev aine omandamine eeldusel, et õppija osaleb sissejuhatavas seminaris, kus tutvustatakse aine sooritamise nõudeid. Seega, hindamisele pääsevad kõik ainele registreeritud tudengid vaatamata sellele, kas ja kui palju on nad õppetöö kontakttundides osalenud. NB! Hindamiseks peab esitama iseseisva töö ja sooritama testi. Nii iseseisvat tööd kui testi võib esitada/sooritada ühe aasta jooksul peale kursuse lõppemist st 2010 aasta sügissemestril toimuva aine sooritusel peavad olema tehtud/esitatud hiljemalt 2012 aasta

	kevadsemestri eelnädalaks välja kuulutatud korduseksami tähtajal/tähtajaks.
Iseseisva töö nõuded	Igal üliõpilasel tuleb kursuse käigus koostada ning kursuse lõpus esitada mahukas arvestuslik iseseisev töö (32 tundi). Iseseisev 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). Iseseisvat tööd võivad õppijad sooritada kuni kolmeliikmelistes gruppides (gruppide koosseis tuleb õppejõuga eelnevalt kooskõlastada).
Eksami hindamiskriteeriumid või arvestuse sooritamiseks vajalik miinimumtase	Hindamiskriteeriumid, millest hindamisel lähtutakse: 1. avatud küsimustega test: küsimused/ülesanded valitakse testi nii, et nad peegeldaks komplektina kolme esimest programmis kirjeldatud õpitulemust; iga küsimus/ülesanne annab teatud arvu punkte; punktid summeeritakse ning hinne kujuneb ülikoolis tavaks kujunenud süsteemi alusel: „A” (90-100%) – silmapaistev ja eriti laiapõhjaline õpiväljundite saavutamise tase, mida iseloomustab väga head taset ületav teadmiste ja oskuste vaba ning loov kasutamine; „B” (80-89%) – väga heal tasemel õpiväljundite saavutamine, mida iseloomustab teadmiste ja oskuste eesmärgipärane ja loov kasutamine. Spetsiifilisemate ja detailsemate teadmiste ja oskuste osas võivad ilmnedä mittesisulised ja mittepõhimõttelised eksimused; „C” (70-79%) – heal tasemel õpiväljundite saavutamine, mida iseloomustab teadmiste ja oskuste eesmärgipärane kasutamine. Spetsiifilisemate ja detailsemate teadmiste ja oskuste osas avaldub ebakindlus ja ebatäpsus; „D” (60-69%) – piisaval tasemel õpiväljundite saavutamine, mida iseloomustab teadmiste ja oskuste kasutamine harjumuspärase olukorras, kuid erandlikes olukordades avalduvad puudujäägid ja ebakindlus; „E” (50-59%) – minimaalselt lubataval tasemel olulisemate õpiväljundite saavutamine, mida iseloomustab teadmiste ja oskuste kasutamine tüüpolekordades piiratud viisidel, kuid erandlikes olukordades avalduvad märgatavad puudujäägid ning ebakindlus; „F” (alla 50%) – õppija on omandanud teadmised ja oskused miinimumtasemest madalamal tasemel. Hinne "F" on negatiivne

	tulemus ning eksam/ test/ vms tuleb uuesti sooritada. 2. iseseisev töö peegeldab kõiki õpitulemusi ning selle hindamine käib vastavalt järgnevatele kriteeriumitele: „+“ kõik täiesti laitmatu (tõstab testil saadud hinnet ühe palli võrra v.a tulemust F ja A) „+v“ põhiosas õiged meetodi valikud ja tõlgendused, mõned vead või puudujäägid (s.h lahendamata ülesandeid, mitte kõige ratsionaalsemaid lahendusi, nappe tõlgendusi jne) (tõstab testil saadud hinnet ühe palli võrra v.a tulemust F ja A) „v+“ pigem õiged meetodi valikud ja tõlgendused, esineb vigu ja/või puudujääke (eksami hindeks jääb testi hinne) „v“ töö põhinõudmisi on järgitud, kuid töös esineb olulisi vigu ning puudujääke (eksami hindeks jääb testi hinne) „v-“ töö põhinõudmisi on järgitud, kuid töös esineb palju olulisi vigu ning puudujääke (langetab testil saadud hinnet ühe palli võrra) „-v“ sobivate lahenduste, kasutatud meetodite jms osakaal alla 50% (töö tuleb ümber teha ja uuesti esitada) „-“ töös ei ole üldse või on vaid mõned üksikud ootuspärased ja õiged lahendused (töö tuleb ümber teha ja uuesti esitada)
Informatsioon kursuse sisu kohta, kursuse jaotumine teemade kaupa sh kontakttundide ajad	Vt eraldi tabel all!

Õppetöö sisu ja ajakava (Tunnid toimuvad kell 16.15-19.45 ruumis T-302)

Kuupäev	Teema, sisu lühikirjeldus
09.09.10	Ülevaade kursuse korraldusest.
Sissejuhatav loeng	Kordav ülevaade kirjeldava statistika põhitõdedest. Sissejuhatus Üldistava statistika põhimõtetesse.
Kodus: videoloeng	Juhuslik valim, selle tähtsus ja moodustamise viisid. Statistiline viga. Valimi-keskväärtuste jaotuse standardhälve e. standardviga.

Lugeda lk 30 -36	Normaaljaotuse idee. Proportsioonid normaaljaotuskõvera all. Väärtuste standardiseerimine
	Üldkogumi keskvaartuse ja teiste arvnäitajate hindamine. Usaldusnivoo. Usaldusintervall. Usalduspiirid.
Kodus: videoloeng	Statistilised olulisustestid - Statistiline hüpotees; selle kontrollimine. Olulisusnivoo ja olulisustõenäosus.
Lugeda lk 37 -51	Kahe valimi keskvaartuste võrdlemine – t-test. Sõltuvad ja sõltumatud valimid. Ühe- ja kahepoolsed testid. I ja II liiki viga
	Mitteparameetrilised meetodid. Risttabelid ja χ^2 -test.
<u>23.09.10</u> Praktikum	Keskvaartuse usaldusintervall, t-test, χ^2 -test
Kodus: videoloeng	Korrata: Statistilised olulisustestid - Statistiline hüpotees; selle kontrollimine. Olulisusnivoo ja olulisustõenäosus.
Lugeda eraldi materjal	Dispersioonanalüüs (ANOVA)
<u>07.10.10</u> Praktikum	ANOVA, Kruskal-Wallise test
Kodus: videoloeng	Statistilise seose liigid. Seosekordajad. Ülevaade korrelatsioonanalüüsist. Korrelatsioonikordaja statistiline olulisus.
Lugeda lk.52-58	Korrelatsioonanalüüs. Korrelatsioonikordaja statistiline olulisus
<u>21.10.10</u> Praktikum	Seoste analüüs
Kodus:	Iseseisev materjali kinnistamine; Iseseisva töö tegemine
<u>04.11.10</u> Praktikum	Praktilised andmeanalüüsi ülesanded kõigi programmis olevate meetodite kasutusega
Kokkuleppel rühmaga	Iseseisvate tööde esimene esitamise tähtaeg!!!
Kokkuleppel rühmaga	Kirjalik test

Õppeainet kureeriv üksus:	Informaatika instituut
Kursuseprogrammi koostaja	Katrin Niglas
Allkiri:	
Kuupäev:	

Kursuseprogramm registreeritud akadeemilises üksuses

Kuupäev	
Õppeassistendi nimi	
Allkiri	