

IFI7041	ANDMEANALÜÜS: statistiline andmestik ja kirjeldav statistika		
Maht: 4EAP	Kontakttundide maht: 42 tundi	Õppesemester: S	Eksam
Eesmärk:	Toetada andmete kogumiseks, töötlemiseks ning lihtsama statistilise analüüsi läbiviimiseks vaja minevate teoreetiliste teadmiste ja praktiliste oskuste omandamist. Tutvustada statistikapaketi kasutusvõimalusi andmete töötlemisel ning statistilisel analüüsil. Arendada teadmiste ja oskuste praktilise rakendamiskogemuse kujunemist, mis võimaldab teha iseseisvalt otsustusi sobiva(te) analüüsimeetodi(te) valikuks ning analüüsi käigus saadud tulemuste korrektseks esitamiseks.		
Aine lühikirjeldus:	Teemad: - Statistiline andmestik, selle kogumine ning töötlemine. - Statistiliste tunnuste tüübid. - Kirjeldav statistika e. andmete kokkuvõtu- ning esitlusmeetodid: erinevad tabelid, diagrammid ja arvnäitajad. - Seoste kirjeldamine: seosekordajad ning risttabelid. - Sobiva analüüsi- või esitlusmeetodi valik. - Mitme valikuvõimalusega küsimuste analüüs. Korraldus: - Kursuse sooritamiseks vajalik töö maht on 104 (4x26) tundi. Sellest kontakttundidena toimuvad seminari tüüpi loengud (12 tundi), praktikumid (28 tundi) ja eksam (2 tundi). - Lisaks tuleb igal üliõpilasel teha mahukas kahest osast koosnev iseseisev töö. Töö teine osa 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). - Hinde saamiseks tuleb sooritada ka kirjalik test. Vt. täpsemalt: „Eksami hindamiskriteeriumid“		
Õpiväljundid:	- Koostab korrektse ülesehitusega andmestiku. - Püstitab andmetest ning eesmärkidest lähtuvalt statistilist analüüsi eeldavaid küsimusi. - Selgitab käsitletud kirjeldava statistika meetodite olemust, teab nende rakendamise tingimusi ning oskab analüüsi tulemusi korrektselt tõlgendada. - Eristab andmete/tunnuste tüüpe ning valib vastavalt andmete tüübile ning andmete kohta esitatud küsimuse sisule sobiva analüüsi meetodi (käsitletud meetodite piires). - Kasutab juhendmaterjali abiga vastavat tarkvara lihtsama andmetöötluse ja -analüüsi läbiviimiseks.		
Hindamismeetodid:	Eksam (hindeline). Hinde kujuneb iseseisva töö ja avatud küsimustega testi koondtulemuste põhjal.		
Õppejõud:	lekt Kairi Osula		
Inglisekeelne nimetus:	Data Analysis: Descriptive Statistics		
Eeldusaine:	Arvutikasutuse elementaarskused IFI6001 mahus		
Kohustuslik kirjandus:	Niglas, K. Videoloengud andmeanalüüsist (saadaval informaatika instituudis) Osula, K. Kursuse materjalid (www.tlu.ee/~kairio/7041) Niglas, K. Statistika loengumaterjale (www.tlu.ee/~katrin/) Niglas, K. (2007) Andmeanalüüs statistikapaketi SPSS 14.00 abil. Põhikursus Tallinn, TLÜ.		
Asenduskirjandus:	Hiob, K. (1995) Matemaatiline statistika. Algkursus koolidele, Tallinn Parring, A.-M., Vähi, M., Käärik, E. (1997) Statistilise andmetöötluse algõpetus, Tartu Tooding, L.M. (2007) Andmete analüüs ja tõlgendamine sotsiaalteadustes, Tartu		

Õppetöös osalemise ja eksamile/arvestusele pääsemise nõuded:	Loengutes osalemine ei ole kohustuslik (eeldab praktikumi ajaks iseseisvat loengumaterjali omandamist). Praktikumides osalemine on kohustuslik. Hindamiseks peab üliõpilane esitama iseseisva töö ja sooritama testi. Nii iseseisvat tööd kui testi võib esitada/sooritada ühe aasta jooksul peale kursuse lõppemist st 2012 aasta sügissemestril toimuva aine sooritusel peavad olema tehtud/esitatud hiljemalt 2014 aasta kevadsemestri eelnädalaks välja kuulutatud korduseksami tähtajal/tähtajaks.
Iseseisva töö nõuded:	Igal üliõpilasel tuleb kursuse käigus koostada kaks iseseisvat tööd. Esimene töö (rühmatöö) koosneb etteantud teemal kaasõpilastele esitluse koostamisest ning ettekandmisest. Teemad ning rühmad jagatakse esimeses loengus. Teine töö (individuaalne 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). Iseseisva töö esitamisel tuleb üliõpilasel oma tööd kaitsta – vastata tööga seonduvatele küsimustele ning põhjendada tehtud valikuid.
Eksami hindamiskriteeriumid	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% Eksami hinde saamiseks peavad olema positiivsele tulemusele esitatud mõlemad osad (kirjalik test, is.töö), st. et ei piisa vaid testi või koduste ülesannete esitamisest. 1. Kirjalik (avatud küsimustega) test moodustab 50% eksamihindest ning seda hinnatakse järgmiste kriteeriumide alusel: Teema: statistiline andmestik, selle kogumine ning töötlemine (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%) Defineerib põhimõisteid; praktiline kasutus- ja seostusoskus puudulik. Nimetab andmete kogumisvõimalusi, valiku põhjendamisel esineb puudujääke. Teema: Kvalitatiivsed ja kvantitatiivsed meetodid uurimustöös. (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%) Eristab kvantitatiivseid ja kvalitatiivseid meetode. 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: 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. 2. Iseseisvad tööd moodustavad eksamihindest 50% ning neid hinnatakse järgmiselt: I töö (rühmatöö) 10% + II töö (individaalne töö) 40% I töö (rühmatöö) ülesandeid hinnatakse skaalal: Väga hea töö (8-10%) – ettekanne vastab teemale ning sellel on selge eesmärk. Teema kohta esitatud info on põhjalik ning korrektne. Hea töö (5-7%) – ettekande sisu on laialivalgus. Esineb üksikuid ebatäpsusi ja materjali võiks olla rohkem. Töö on arvestamata (0-4%) II töö (individaalse töö) ülesandeid hinnatakse skaalal: Väga hea töö (34-40%) - 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 (27-33%) - 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 (21-26%) - 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 (0-20%)
--	--

Informatsioon kursuse sisu kohta (kursuse jaotumine teemade kaupa sh kontakttundide ajad). **II rühm.**

Kuupäev	Teema, sisu lühikirjeldus
1. LOENG 5.09.2012 (12.15-13.45) T-416	Sissejuhatus ainesse. Ülevaade kursuse korraldusest. Empiiriliste uuringute põhietapid. Kvantitatiivsed ja kvalitatiivsed meetodid uurimistöös. Andmete kogumine ja analüüsiks ettevalmistamine. Statistilised andmed, tunnusetüübid.
Praktikum 5.09.2012 (14.15-15.45) T-302	Küsitluse koostamine. Tunnusetüübid. Andmekogumismeetodid. Internetiküsitluse koostamise erinevad vahendid.
Praktikum 12.09.2011 (14.15-15.45) T-302	Programmi PASW Statistics 18.0 – ülevaade. Andmetabeli loomine. Tunnuste defineerimine. Andmete sisestamine ja redigeerimine.
2. LOENG 19.09.2012 (12.15-13.45) T-416	Esmane ülevaade andmetest. Sagedustabelid. Andmete esitamise erinevad viisid.
Praktikum 19.09.2012 (14.15-15.45) T-302	Andmefailide transport SPSSi ja teiste programmide vahel. Enamlevinumad andme- ja failiteisendused: tunnuste ümberkodeerimine, uute tunnuste loomine, väärtuste järjestamine, tekstitunnuse muutmine numbriliseks, väärtuste selekteerimine ja andmestiku osadeks jagamine.
Praktikum 26.09.2012 (14.15-15.45) T-302	Sagedustabelid. PASW tulemileht. Tabelite kujundamine. (Kommenteeriva) teksti lisamine tulemustelehele.
3. LOENG 3.10.2012 (12.15-13.45) T-416	Andmete graafiline kirjeldamine.
Praktikum 3.10.2012 (14.15-15.45) T-302	Andmete graafiline kirjeldamine. Sektordiagramm, tulpdiagramm.
Praktikum 10.10.2012 (14.15-15.45) T-302	Andmete graafiline kirjeldamine. Histogramm. Interaktiivsed diagrammid SPSSis.
4. LOENG	Kirjeldavad arvnäitajad. Keskmist tendentsi väljendavad arvnäitajad. Hajuvust

17.10.2012 (12.15-13.45) T-416	väljendavad arvnäitajad. Jaotuse kuju. Asümmeetrilised jaotused. Bimodaalne jaotus. Normaaljaotuse idee. Proportsioonid normaaljaotuskõvera all.
Praktikum 5.09.2012 (14.15-15.45) T-302	Kirjeldavad arvnäitajad – nende arvutamine ja tõlgendamine.
22.10.2012-28.10.2012	Iseseisva töö nädal
Praktikum 31.10.2012 (14.15-15.45) T-302	Andmestiku kirjeldamine, kasutades õpitud meetode.
5. LOENG 7.11.2012 (12.15-13.45) T-416	Erinevuste uurimine e gruppide võrdlemine. Korrelatsioonanalüüs.
Praktikum 7.11.2012 (14.15-15.45) T-302	Mitmemõõtmelised sagedustabelid. Keerulisemad diagrammid. Võrdlevad diagrammid.
Praktikum 14.11.2012 (14.15-15.45) T-302	Risttabelid. Korrelatsioonanalüüs.
6. LOENG 21.11.2012 (12.15-13.45) T-416	Korrelatsioonanalüüs. Üldine kordamine.
Praktikum 21.11.2012 (14.15-15.45) T-302	Korrelatsioonanalüüs. MPAR kordajad. Andmestiku reliaabluse kontroll.
Praktikum 28.11.2012 (14.15-15.45) T-302	Üldine kordamine: kõikide õpitud meetodite kasutamine.
7. LOENG 5.12.2012 (12.15-13.45) T-416	Kirjalik test (esimene võimalus).
Praktikum 5.12.2012 (14.15-15.45) T-302	Iseseisva töö ülesannete lahendamine ja kaitsmine.
Praktikum 12.12.2012 (14.15-15.45) T-302	Iseseisva töö ülesannete kaitsmine.
9.01.2013 (12.15-13.45)	Kirjalik test (teine võimalus)
9.01.2013 (14.15-15.45)	Iseseisva töö ülesannete kaitsmine (kokkuleppel õppejõuga)

Õppeainet kureeriv üksus:	Informaatika instituut
Kursuseprogrammi koostaja:	Kairi Osula
Allkiri:	
Kuupäev:	18.08.2012