

# Statistiline andmeteadus ja visualiseerimine

[andmeteadus.github.io/2016](https://andmeteadus.github.io/2016)

Mark Gimbutas, Siim Viigand

[markgimbutas@gmail.com](mailto:markgimbutas@gmail.com)

# Õppetöö korraldus

- praktikumid
  - kokku 9 tk
  - kolmapäeviti 12:15 ruumis 203
  - küllalt soovituslik teha oma laptopiga
- loenguid ei ole
- kodutööd
  - kokku 8 tk
- projekt
- kontaktaadress [markgimbutas@gmail.com](mailto:markgimbutas@gmail.com)

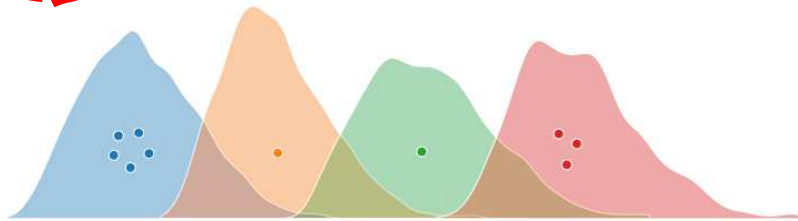
## Aine läbimiseks on vaja

1. **Saada vähemalt 150 punkti** kodutööde eest.
  - Iga kodutöö kuni 25 punkti, kokku kuni  $8 \times 25 = 200$  punkti.
  - Kodutöodes on boonusülesandeid, mis annavad lisapunkte.
2. **Teha projekt**, kirjutada selle kohta populaarteaduslikus stiilis ülevaade ja saadud tulemusi suuliselt esitleda.
  - Kas üksinda või kahepeale.
  - Teema valik: otsige ise või oodake, kuni meie midagi pakume.
  - Teemad, tähtajad jms ilmuvad lehele  
<http://andmeteadus.github.io/2016/projektid/>

# Kodutööd

Statistiline andmeteadus ja visualiseerimine MTMS.01.100

Praktikumid Projektid



## Õppeaine lühikirjeldus

Tegemist on rakendusliku statistikakursusega, kus:

- Probleempüstitusena kasutatakse huvitavaid päriselulisi andmestikke.
- Rakendatakse sobivaid statistilisi meetodeid, et püstitatud küsimustele vastust leida.
- Saadud tulemuste teistele kommunikeerimiseks illustreeritakse neid sobivate joonistega.

Järgnev skeem võtab lühidalt kokku õppeaine olemuse.

## Praktikumimaterjalid

### Praktikum 1

10.02.2015

Sissejuhatus ainesse, reprodutseeritavasse analüüsi ja automatiseeritud statistikasse. Tarkvara R meeldetuletamine. Pakett knitr. Markdown. Pakett dplyr. Näide: automaatne aktsiaturu ülevaade.

- Kursuse “Rakendustarkvara: R” õppematerjal.
- Praktikumis kasutatud materjalid:
  - Slaidid (korralduslik info)
  - Logige [Socrative](#) tagasiside keskkonda koodiga 5fc6f3ec
  - knitr ja R Markdown sissejuhatus
  - dplyr minimalistlik näide
- Automaatne aktsiaturu ülevaade
- Kodutöö
  - Tähtaeg: 16.02.2015 kell 23.59 GMT

- tähtaeg järgneval teisipäeval kell 23:59
- saata aadressile [markgimbutas@gmail.com](mailto:markgimbutas@gmail.com)

# Projekt

- Üksinda või kahepeale tuleb analüüsida ja visualiseerida **mingid andmed**.
  - Teema valimise tähtaeg on 15.04.2016.
  - Teema (s.t. andmed) võib **otsida ise** või oodata **meie pakkumist** (tulevad 01.04.2016).
  - Soovitame ise otsimisega tegeleda alles pärast 5. praktikumi (andmete kraapimine veebist, 16.03.2016).
- Tudengite eelmisel aastal tehtud projektid on <http://andmeteadus.github.io/2015/projektid/> .
  - Suur ideaal, mille poole projektides püüelda, on [www.stat24.ee](http://www.stat24.ee) .
- Projekti juhend [http://andmeteadus.github.io/2016/projekt\\_juhend/](http://andmeteadus.github.io/2016/projekt_juhend/) (hetkel on tähtajad vanad).

# Aeg

- 3 EAP = 78 h jagu töömahtu jaguneb
  - 9 x 1,5 h = 13,5 h praktikumid
  - 8 x 5 h = 40 h kodutööd
  - 23 h projekti koostamine
  - 1,5 h projekti esitlemine (ja teiste esitluste kuulamine)



MTMS.01.100

teoreetilised  
statistika-  
teadmised

tegelikkuses  
vajalikud  
oskused