

Kuidas saada vähese vaevaga võimalikult häid hindeid?

Risto Hinno

Friday, June 26, 2015

Kuidas saada vähese vaevaga võimalikult häid hindeid?

Sissejuhatus

See oli juba kolmas järjestikune öö, mil tudeng Keit “Ogar” Agarilt viis une õudusunenägu. “Te olite võrdsed kandidaadid, kuid parema keskmise hinde tõttu võtsime tööle teise kandidaadi,” ütles OÜ Tunnetuslik Homöopaatia personalijuht. Töö sai Keiti alumine naaber, kes õpib 15. sajandi inglise kirjandust. Keit arvab, et seegi unenägu võib olla täidminev ning tal on vaja head vastuargumenti personalijuhile. Äkki on mõnes teaduskonnas kergem paremat hinnet saada, mõtles Keit.

Ta teab, et finantsseisu parandamiseks ei saa lootma jääda vaid tööintervjuule. Mida teeks sellises olukorras tema ületee naaber, kes on saanud mitmeid EL toetusi ja sõidab kullatud istmesoojendusnupuga varustatud Mercedes-Benz maasturiga? Ilmselt uuriks, kuidas võimalikult vähese vaevaga parandada oma keskmist hinnet, et nii stipendiumit saada.

Kuna kolm plaani on ikka parem kui kaks, mõtles Keit, mida võtaks ette tema ettevõtlik tuttav, kes juhib alkoholilaenutuse ettevõtet. Pärast pingsat mõtlemist süttis Keiti peas lambike: ehk teeks ta abistava rakenduse kasvavale tudengipopulatsioonile, kelle eesmärk on võimalikult lihtsalt diplom kätte saada?

Selles praktikumis uurimegi:

- kas tõesti on mõnes teaduskonnas kergem paremat hinnet saada,
- kas on aineid, milles on märgatavalt kergem paremat hinnet saada,
- kuidas teha veebirakendusi R-is.

Tutvumine andmestikuga

Laadi ÖISist alla andmestik `UT_13_14_hinded.csv`. Selles andmestikus on 2013/2014 õppeaasta kõigi TÜ õppeainete kohta, mis olid eristava hindamisega, toodud üliõpilaste õppetulemuste jaotus. Lisaks on teada

- semester (sügis/kevad)

-teaduskond (TÜ struktuuriüksuse lühend)

- AR Arstiteaduskond
- EC Euroopa kolledž
- FL Filosoofiateaduskond
- KK Kehakultuuriteaduskond
- LO Loodus- ja tehnoloogiateaduskond
- MJ Majandusteaduskond
- MT Matemaatika-informaatikateaduskond
- NC Narva kolledž

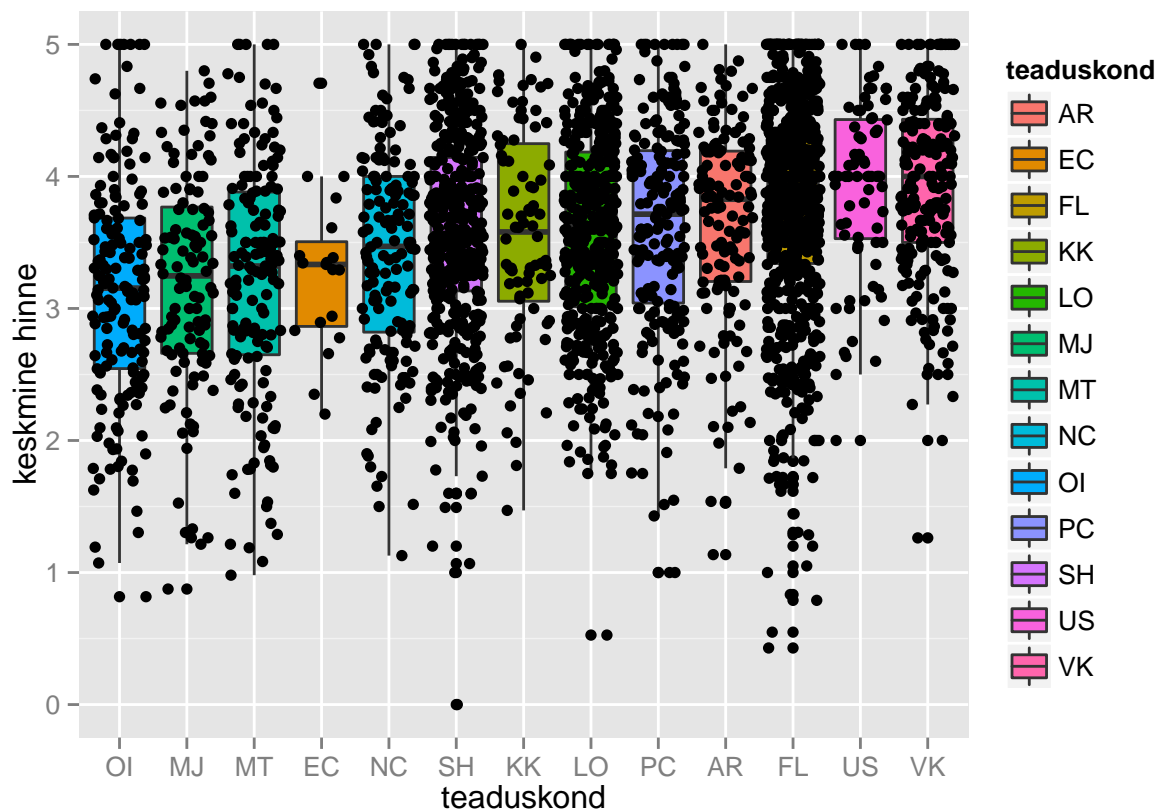
- OI Õigusteaduskond
- OO Õppeosakond
- PC Pärnu kolledž
- SH Sotsiaal- ja haridusteaduskond
- US Usuteaduskond
- VK Viljandi kultuuriakadeemia
- aine kood
- aine nimi
- ainele registreerunud üliõpilaste arv
- õppeainele registreerunud üliõpilaste õppetulemuste (A, B, C, D, E, F, mitte ilmunud) jaotus

Ülesanne 1 (4 punkti)

Kas mõnes teaduskonnas saavad üliõpilased paremaid hindeid kui mõnes teises? Kuidas seda visuaalselt kontrollida? Tee seda.

```
hinded=read.csv2("./data/UT_13_14_hinded.csv")
library(dplyr)
hinded2=mutate(hinded,
               keskmine = (A*5+B*4+C*3+D*2+E*1)/(registreerunuid-mi))
#osades on hindeid saanud rohkem kui ainele reganud, eemaldan need vaatlused
hinded2_sub=subset(hinded2, keskmine<=5& keskmine>=0)
#leiame järjekorra, keskmiste kaupa
hinded2_order=with(hinded2_sub, reorder(teaduskond,keskmine, median))

#teeme ploti
library(ggplot2)
ggplot(hinded2_sub, aes(x=hinded2_order, y=keskmine))+
  geom_boxplot(aes(fill=teaduskond))+
  geom_jitter()+
  xlab("teaduskond")+
  ylab("keskmine hinne")
```

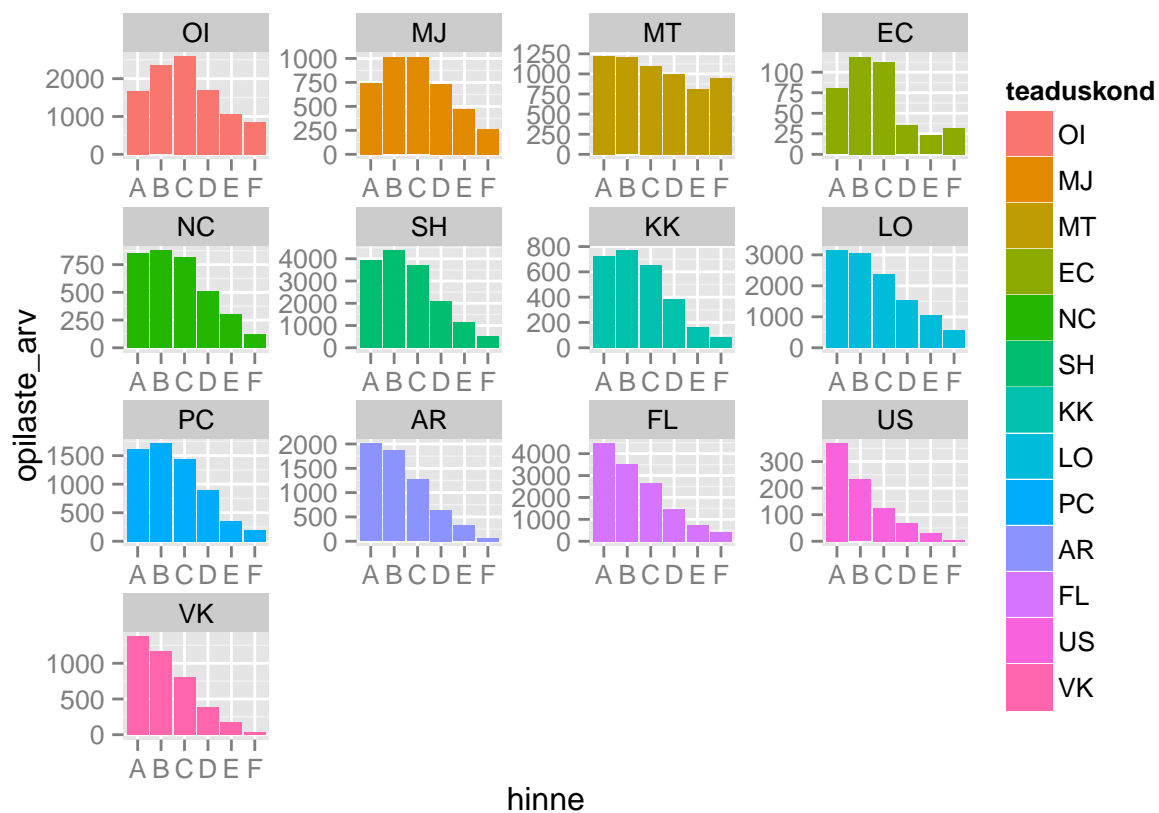


```
#teise ploti jaoks teeme andmed lühikeseks
library(reshape2)
hinded2$mi=NULL
hinded2_melt=melt(hinded2, measure.vars = c("A", "B", "C", "D", "E", "F"))

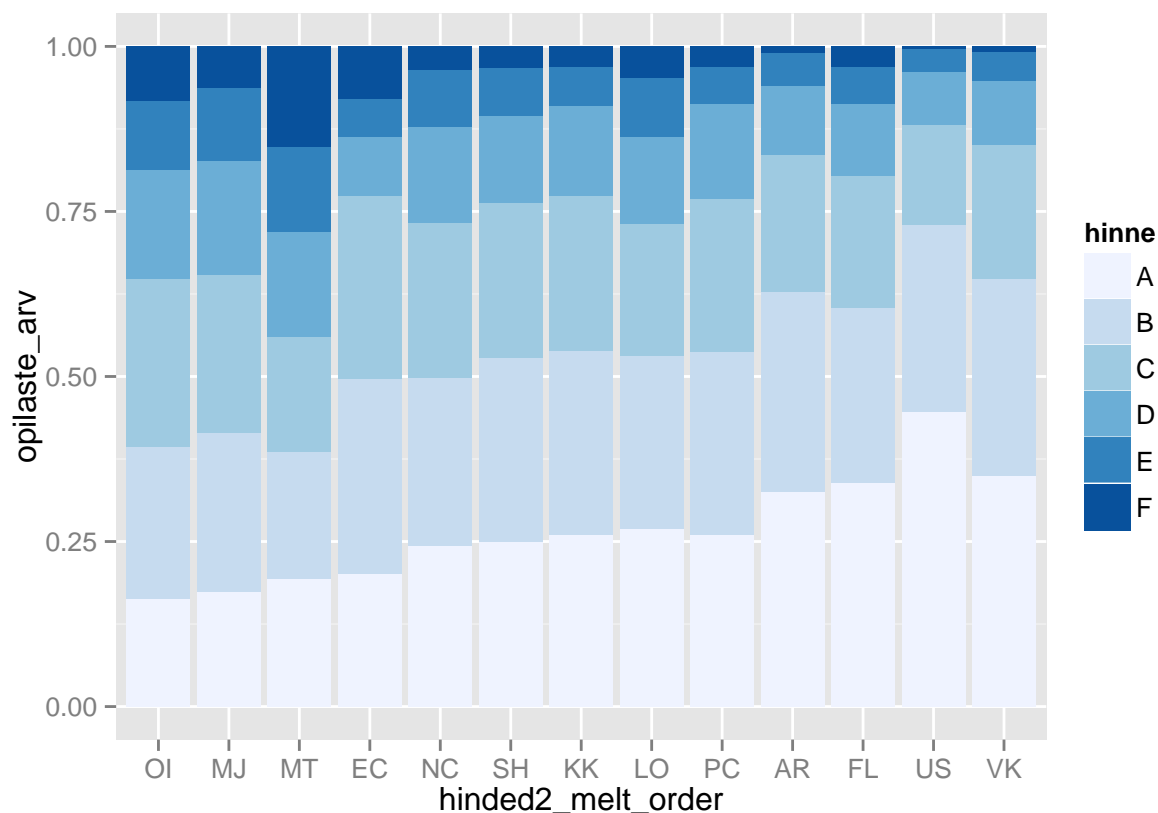
names(hinded2_melt)[7:8]=c("hinne", "opilaste_arv")

#leiam järjekorra, keskmiste kaupa
hinded2_melt_order=with(hinded2_melt, reorder(teaduskond,keskmine, median))

#leiam õige teaduskonna järjestuse keskmise hinde järgi
levelid=levels(hinded2_melt_order)
hinded2_melt2=hinded2_melt
#paneme selle tabelisse nii, et graafikul kuvaks õiges järjekorras
hinded2_melt2$teaduskond <- factor(hinded2_melt2$teaduskond, levels = levelid)
#teeme ploti
ggplot(hinded2_melt2, aes(x=hinne, y=opilaste_arv, fill=teaduskond))+
  geom_bar(stat="identity")+
  facet_wrap(~teaduskond, scales = "free")
```



```
#teeme kolmanda ploti
ggplot(hinded2_melt, aes(x=hinded2_melt_order, y=opilaste_arv, fill=hinne))+
  geom_bar(position="fill", stat="identity")+
  scale_fill_brewer()
```



Ülesanne 2 (4 punkti)

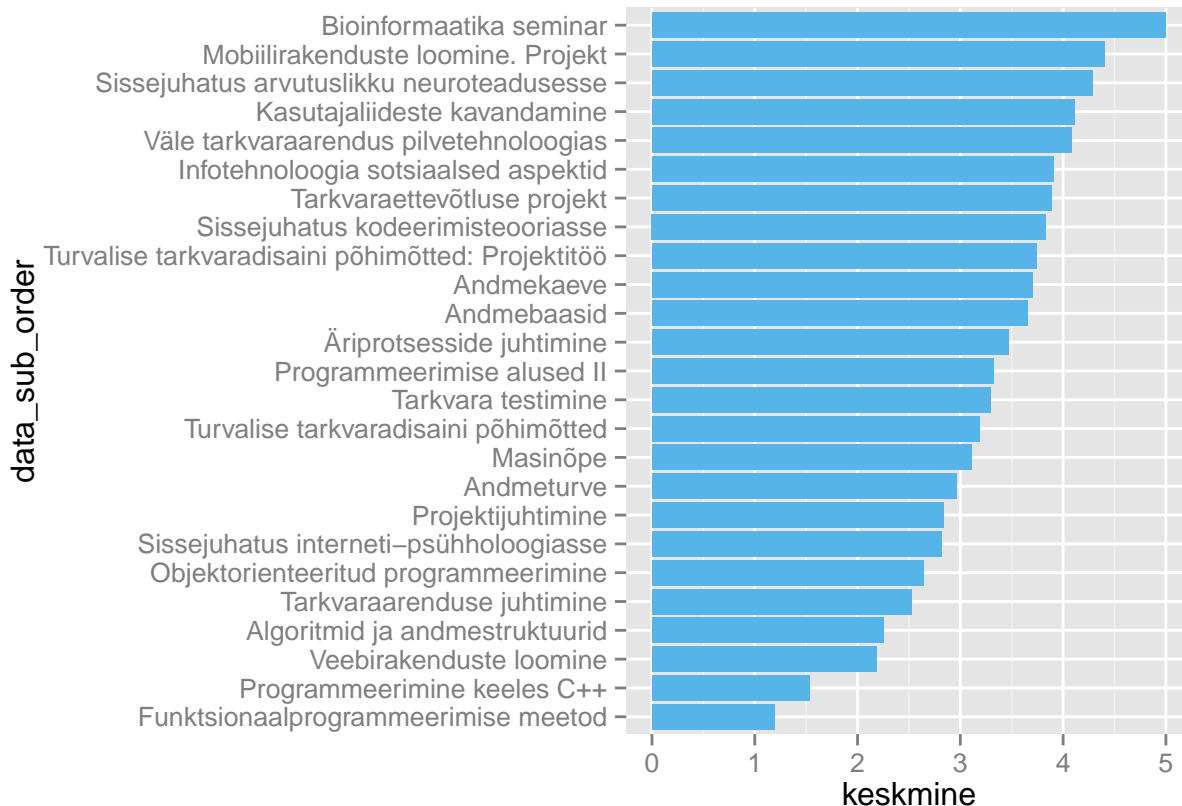
Kas mõnes õppeaines saavad üliõpilased paremaid hindeid kui mõnes teises? Millistes ainetes täpsemalt? Visualiseeri.

```
data=hinded
sem=c("K")
teaduskond1=c("MT")

#teeme subseti andmetest
data_subset=subset(data, teaduskond==teaduskond1 & semester==sem)
library(dplyr)
data_subset=mutate(data_subset,
  keskmine = (A*5+B*4+C*3+D*2+E*1)/(registreerunuid-mi))
#osades on hindeid saanud rohkem kui ainele reganud, eemaldan need vaatlused
data_subset=subset(data_subset, keskmine<=5 & keskmine>=0)
#hoiame alles 25 ainet
data_subset=data_subset[1:25,]
#leiame ainete järjekorra, keskmiste kaupa
data_sub_order=with(data_subset, reorder(aine,keskmine, median))

#teeme ploti
library(ggplot2)
ggplot(data_subset, aes(y=keskmine, x=data_sub_order))+
```

```
geom_bar(stat="identity", fill="#56B4E9")+
coord_flip()
```



```
#sama asi funktsioonina
plot_oppeained = function(data, tk, sem){
  library(dplyr)
  data_subset=data %>%
    subset(teaduskond==tk & semester==sem) %>%
    mutate(keskmine = (A*5+B*4+C*3+D*2+E*1)/(registreerunuid-mi)) %>%
    subset(keskmine<=5& keskmine>=0)

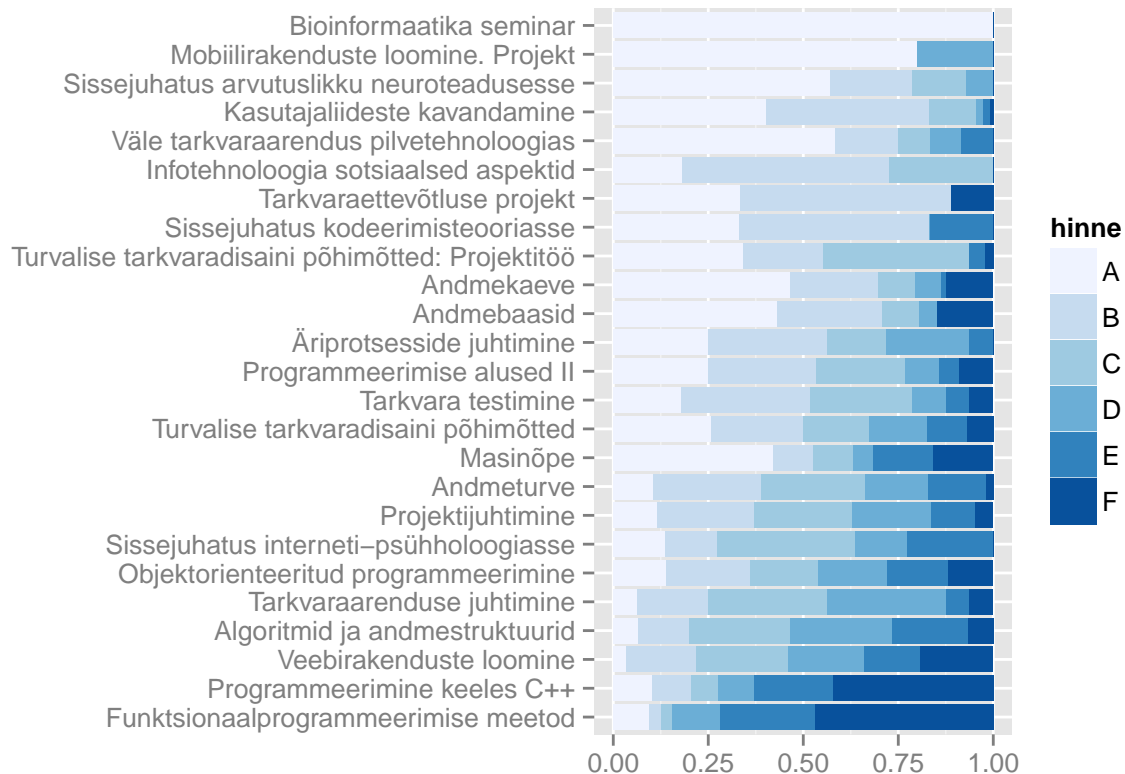
  #hoimae ianult 25 rida alles
  data_subset=data_subset[1:25,]
  #teeme kitsaks andmetabeli
  library(reshape2)
  data_subset$mi=NULL
  data_subset_melt=melt(data_subset,
                        measure.vars = c("A", "B", "C", "D", "E", "F"))
  names(data_subset_melt)[7:8]=c("hinne", "opilaste_arv")
  #teeme ploti
  library(ggplot2)
  ggplot(data_subset_melt, aes(x=reorder(aine,keskmine, median),
                                y=opilaste_arv,fill=hinne))+
    geom_bar(position="fill", stat="identity")+
    scale_fill_brewer()+
    xlab("")+
    ylab("")}
```

```

    ylab("")+
    coord_flip()
}

plot_oppeained(data=hinded, tk="MT", sem="K")

```



Boonusülesanne 2: Kuidas optimeerida õppeaine tagasisidet?

Õppejõud Peeter “Bijeksioon” Ratsionaalil kästi õppeaine ümber korraldada. Tema õpetatavate ainete tagasiside on siiani olnud kehv, kuid ta on veendunud, et see pole mitte viletsa õpetamise ja halva ainekorralduse tõttu, vaid et üliõpilased annavadki paremat tagasisidet sellistele ainetele, kus eriti midagi tegema ei pea ning kus saab vähese vaevaga häid hindeid.

Uri, kas Peetri väited peavad paika. See tähendab, et kas tõepoolest on tal lootust saada ainele paremat tagasisidet, kui selle töömaht on väiksem, kui ette nähtud, ning kui tudengitele panna paremaid hindeid.

Laadi ÕISist alla andmestik UT_13_14_tagasiside.csv. (Praktikumis kasutatud andmestik oli selle alamandmestik.) Siin on toodud 2013/2014 õppeaasta kõigi TÕ õppeainete kohta (st nii eristava kui ka mitteeristava hindamisega):

- semester (sügis/kevad)
- teaduskond (TÕ struktuuriüksuse lühend, samasugused, nagu praktikumis kasutatud andmestikus)
- aine kood

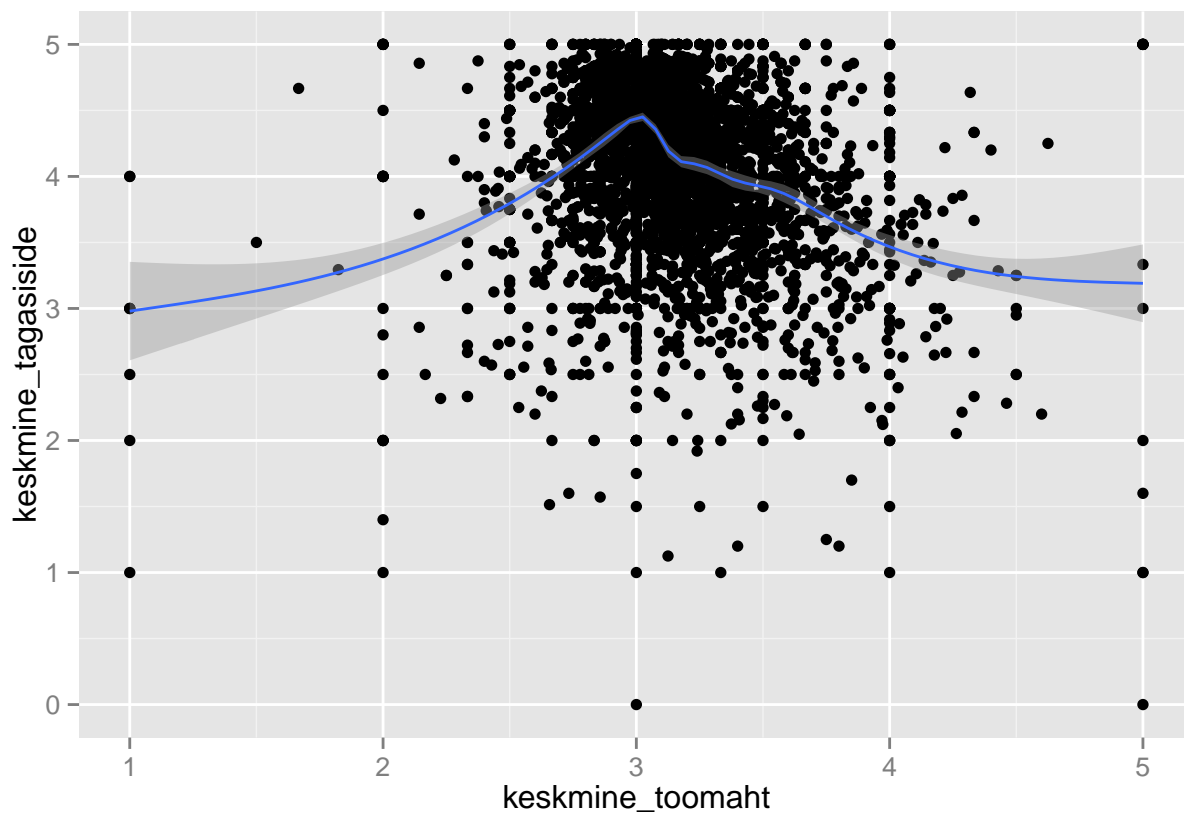
- aine nimi
- ainele registreerunud üliõpilaste arv
- õppeaine töömahtu puudutava tagasiside küsimuse vastuste jaotus
- tagasiside_1 - aine töömaht oli tunduvalt väiksem
- tagasiside_2 - aine töömaht oli mõnevõrra väiksem
- tagasiside_3 - aine töömaht oli selline, nagu ette nähtud
- tagasiside_4 - aine töömaht oli mõnevõrra suurem
- tagasiside_5 - aine töömaht oli tunduvalt suurem
- tagasiside küsimuse “Kokkuvõttes hindan antud õppeainet hindega” tulemuste jaotus
- veerud tagasiside_A, ..., tagasiside_F
- õppeainele registreerunud üliõpilaste õppetulemuste jaotus
- eristava hindamisega ainete jaoks veerud hinne_A, ..., hinne_mi
- mitteeristava hindamisega ainete jaoks veerud hinne_arv, hinne_marv, hinne_mi

Ülesanded:

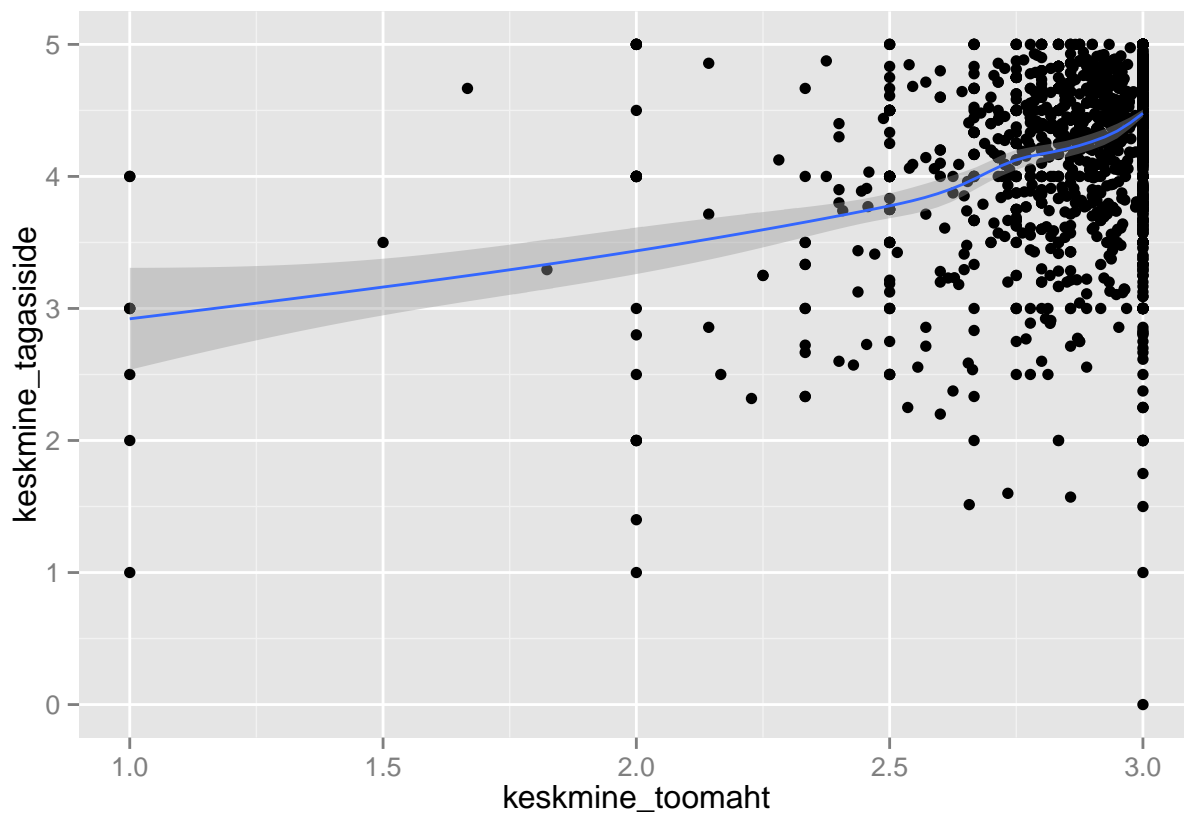
(3 boonuspunkti) Uuri, kas õppeaine tagasiside skoor üleüldse sõltub aine töömahust, ning kas ette nähtust väiksem töömaht tagab ainele parema tagasiside. Visualiseeri.

```
tagasiside=read.csv2("./data/UT_13_14_tagasiside.csv")
#tagasiside$hinne_arv=NULL
#tagasiside$hinne_marv=NULL
#tagasiside$hinne_mi=NULL
tagasiside=tagasiside %>%
  mutate(keskmine_hinne = (hinne_A*5+hinne_B*4+hinne_C*3+hinne_D*2+hinne_E*1+
    hinne_F*0)/
    (hinne_A+hinne_B+hinne_C+hinne_D+hinne_E))%>%
  mutate(keskmine_toomaht = (toomaht_1*1+toomaht_2*2+toomaht_3*3+toomaht_4*4+
    toomaht_5*5)/
    (toomaht_1+toomaht_2+toomaht_3+toomaht_4+
    toomaht_5))%>%
  mutate(keskmine_tagasiside = (tagasiside_A*5+tagasiside_B*4+tagasiside_C*3+
    tagasiside_D*2+tagasiside_E*1+
    tagasiside_F*0)/
    (tagasiside_A+tagasiside_B+tagasiside_C+
    tagasiside_D+tagasiside_E+
    tagasiside_F))

#seos keskmise tagasiside ja töömaht
ggplot(tagasiside, aes(x=keskmine_toomaht, y=keskmine_tagasiside))+
  geom_point()+
  geom_smooth()
```

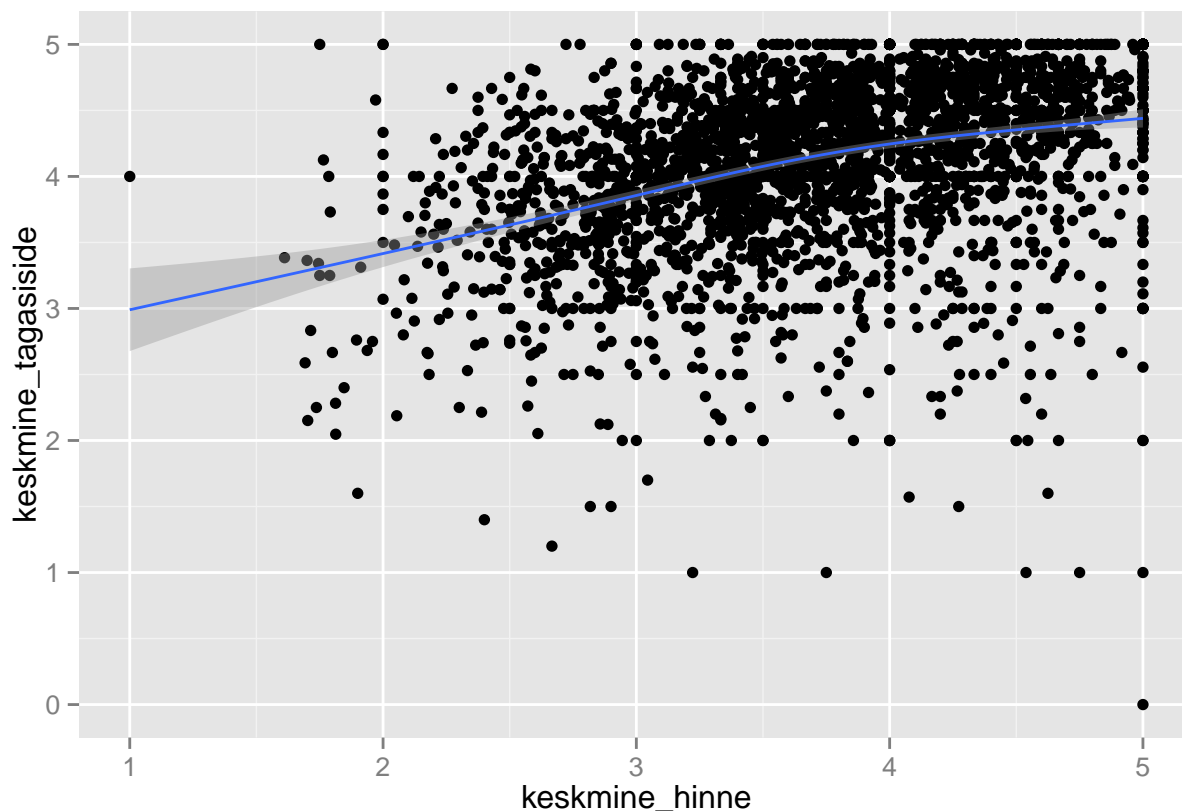
```
#kas ette nähtust väiksem töömaht tagab ainele parema tagasiside  
ggplot(subset(tagasiside, keskmine_toomaht<=3), aes(x=keskmine_toomaht,  
                                                    y=keskmine_tagasiside))+  
  geom_point()+  
  geom_smooth()
```



Jooniselt on näha, et keskmisest väiksema töömahu korral on keskmine tagasiside hoopis madalam, kui optimaale töökoormuse korral.

(3 boonuspunkti) Uuri, kas õppeainele antud tagasiside üldse sõltub tudengite õppetulemustest seal aines, ning milline see sõltuvus on. Visualiseeri.

```
#kas tagasiside sõltub hindest
ggplot(tagasiside, aes(x=keskmine_hinne, y=keskmine_tagasiside))+
  geom_point()+
  geom_smooth()
```



```
#teeme regressioonimudeli
regressioon=lm(data=tagasiside, keskmine_hinne ~ keskmine_tagasiside)
summary(regressioon)
```

```
##
## Call:
## lm(formula = keskmine_hinne ~ keskmine_tagasiside, data = tagasiside)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.70144 -0.45144 -0.00125  0.45967  2.89543
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.10457    0.07781   27.05  <2e-16 ***
## keskmine_tagasiside 0.39922    0.01865   21.41  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6617 on 2953 degrees of freedom
## (1986 observations deleted due to missingness)
## Multiple R-squared:  0.1343, Adjusted R-squared:  0.134
## F-statistic: 458.2 on 1 and 2953 DF,  p-value: < 2.2e-16
```

Jooniselt on näha, et mida kõrgem on keskmine hinne, seda kõrgem on ka tagasiside.