

# KOV valimised

*Risto Hinno*

*Friday, June 26, 2015*

## 2013 KOV valimiste hääletustulemuste andmestik

Loe sisse andmestik ja tutvu andmetega. Andmestikus on järgnevad tunnused:

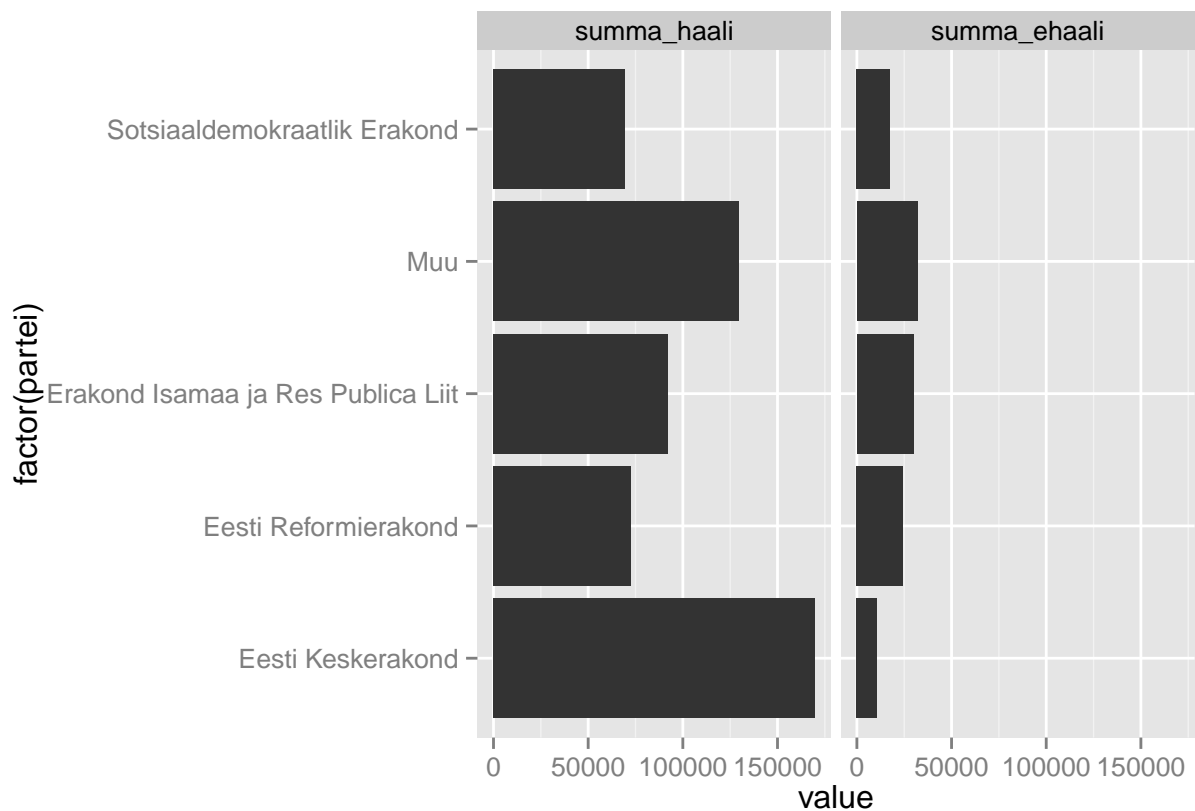
- id: kandidaadi ID
- KOV
- maakond
- nimekiri - partei või valimisliit, kuhu kandidaat kuulub
- partei - kas Keskerakond, Reformierakond, IRL, SDE või Muu
- nr - kandidaadi number
- haali - kandidaadi poolt saadud koguhääle arv (st paberhääled + ehääled)
- ehaali - kandidaadi poolt saadud e-hääle arv

```
valimised=read.csv2("./data/KOV_valimised_2013.csv")
```

## Ülesanne 3 (2 punkti)

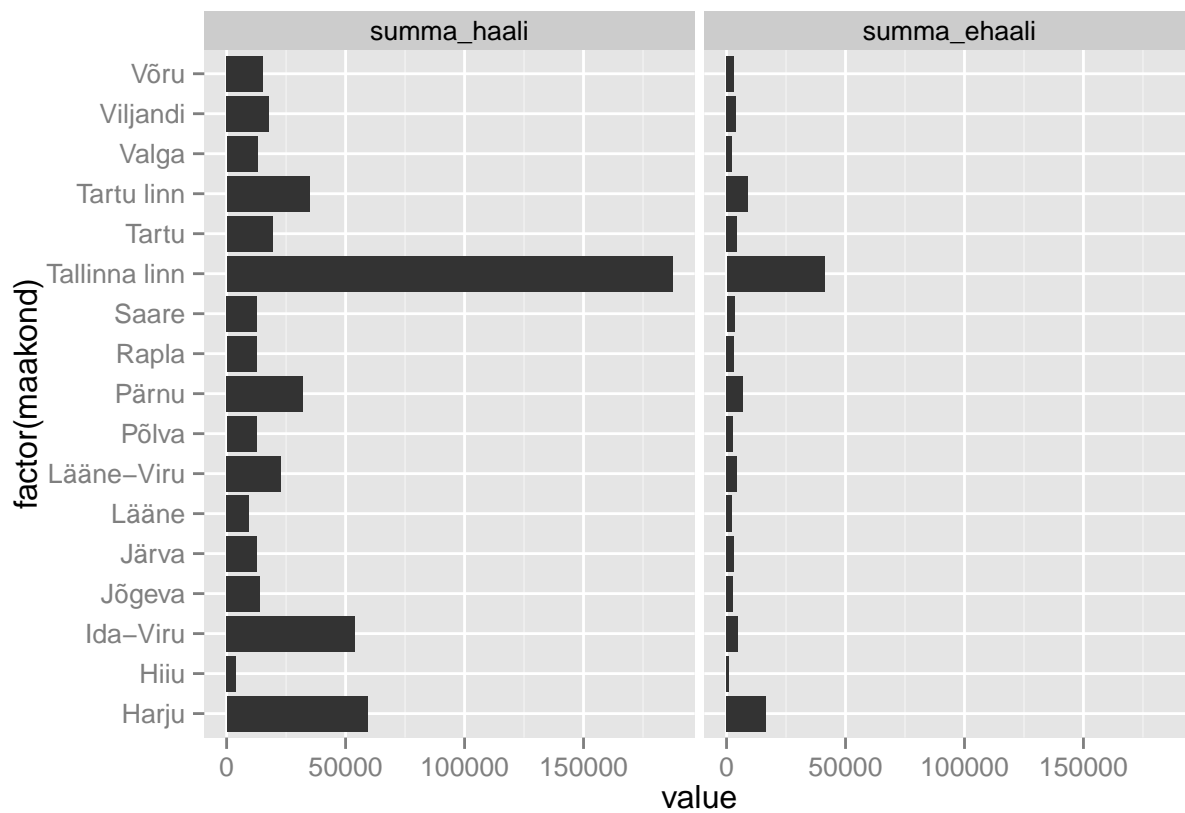
Tee 3 joonist, mis iseloomustavad hästi seda andmestikku.

```
library(dplyr)
library(reshape2)
partei_group=valimised%>%
  group_by(partei)%>%
  summarise(summa_haali=sum(haali),
            summa_ehaali=sum(ehaali))%>%
  melt()
#üldine hääle ja e-jääle jagunemine parteide vahel
library(ggplot2)
ggplot(partei_group, aes(x=factor(partei), y=value))+
  geom_bar(stat = "identity")+
  facet_wrap(~variable)+
  coord_flip()
```

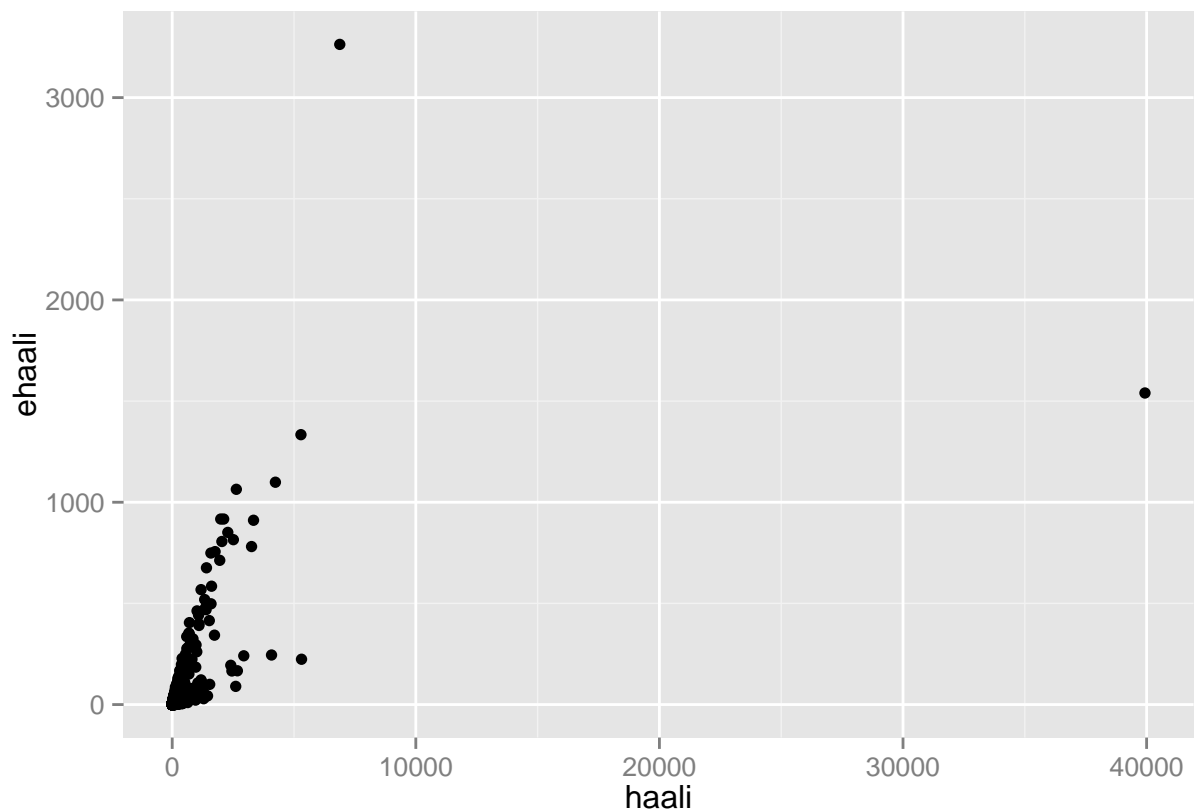


```
##üldine häälte ja e-jäälte jagunemine maakondade vahel
maakond_group=valimised%>%
  group_by(maakond)%>%
  summarise(summa_haali=sum(haali),
            summa_ehaali=sum(ehaali))%>%
  melt()

ggplot(maakond_group, aes(x=factor(maakond), y=value))+
  geom_bar(stat = "identity")+
  facet_wrap(~variable)+
  coord_flip()
```



```
#e-hääslte ja hääslte arvu vaheline seos
ggplot(valimised, aes(x=haali, y=ehaali))+
  geom_point()
```



## Ülesanne 4 (2 punkti)

Kontrolli visuaalselt Benfordi seaduse kehtimist:

- e-hääle arvul,
- paberhääle arvul,
- koguhääle arvul.

Seejärel tee eelnevad joonised ka erakondade kaupa. Kommenteeri tulemusi.

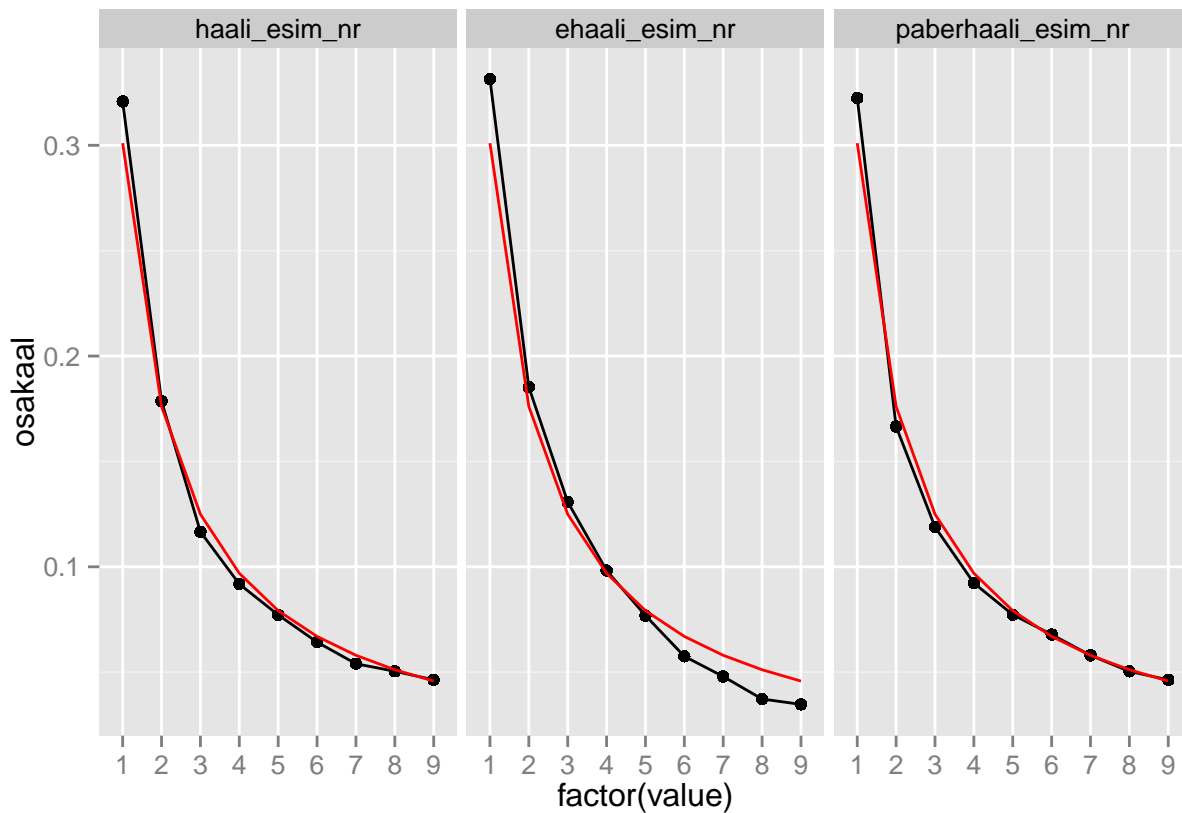
```
#arvutame paberhääle arvu
valimised$paberhaali=valimised$haali-valimised$ehaali
#lõikame kõigi esimese nri
valimised$haali_esim_nr=substr(valimised$haali, 1, 1)
valimised$ehaali_esim_nr=substr(valimised$ehaali, 1, 1)
valimised$paberhaali_esim_nr=substr(valimised$paberhaali, 1, 1)

#teeme benfordi jaotusele vastava jaotuse
library(BenfordTests)
benford=as.data.frame(pbenf(digits = 1))
#arvutame iga esinumbri counti
library(reshape2)
valimised2=melt(valimised[, c("partei", "haali_esim_nr", "ehaali_esim_nr",
```

```

                                "paberhaali_esim_nr"]],id=c("partei"))
valimised2=subset(valimised2, value!=0)
#arvutame iga grupi (hääle liigi) counti, iga hääli liigi iga esimese nr counti
#ja iga esinumbri osakaalu hääle liigis
library(dplyr)
valimised_liik=valimised2 %>%
  group_by(variable) %>%
  mutate(kokku=n()) %>%
  group_by(variable, value) %>%
  mutate(kokku_grupp=n(),
         osakaal=kokku_grupp/kokku)
#teeme ploti
ggplot(valimised_liik,aes(x=factor(value), y=osakaal, group=1))+
  geom_point()+
  geom_line()+
  geom_line(data=benford, aes(x=Var1, y=Freq,group=1), colour="red")+
  facet_wrap(~variable)

```



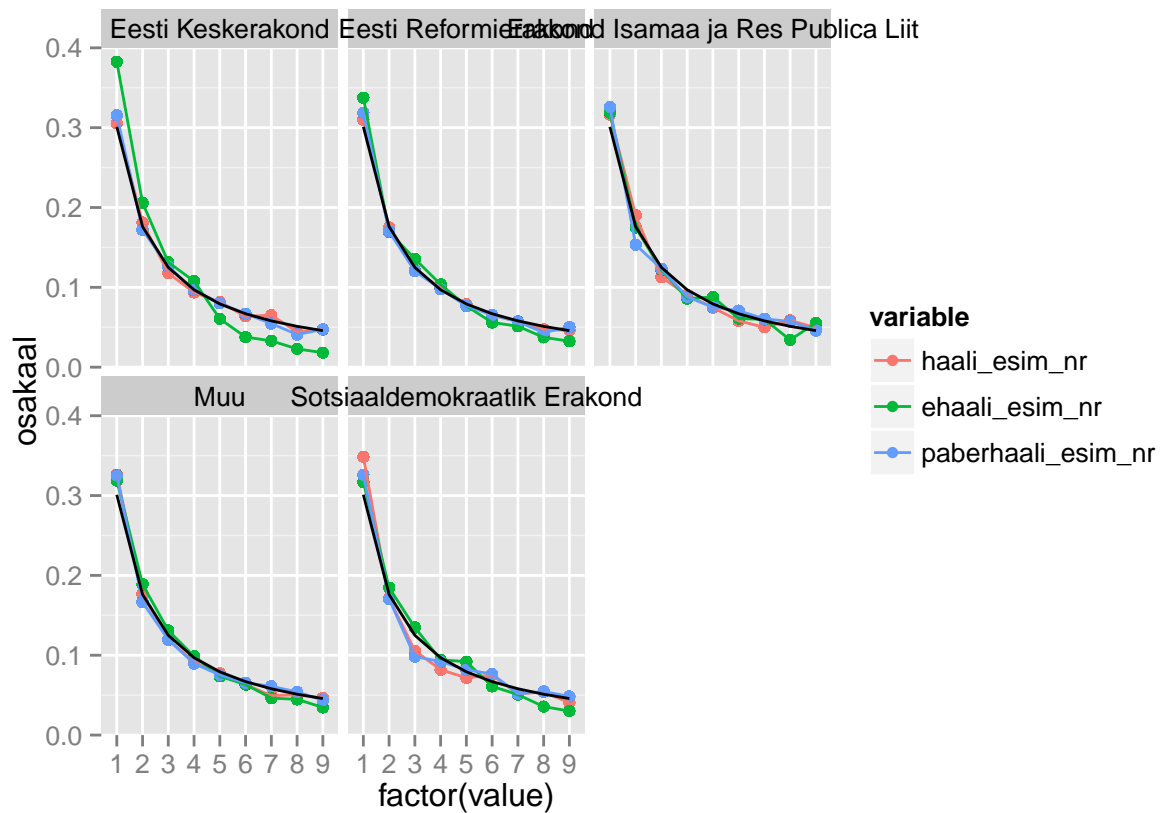
```

#teeme sama asja erakondade kaupa
valimised_liik_partei=valimised2 %>%
  group_by(partei, variable) %>%
  mutate(kokku=n()) %>%
  group_by(partei, variable, value) %>%
  mutate(kokku_grupp=n(),
         osakaal=kokku_grupp/kokku)

```

```
#teeme ploti
ggplot(valimised_liik_partei,aes(x=factor(value), y=osakaal, group=variable,
                                colour=variable))+

  geom_point()+
  geom_line()+
  geom_line(data=benford, aes(x=Var1, y=Freq,group=1), colour="black")+
  facet_wrap(~partei)
```

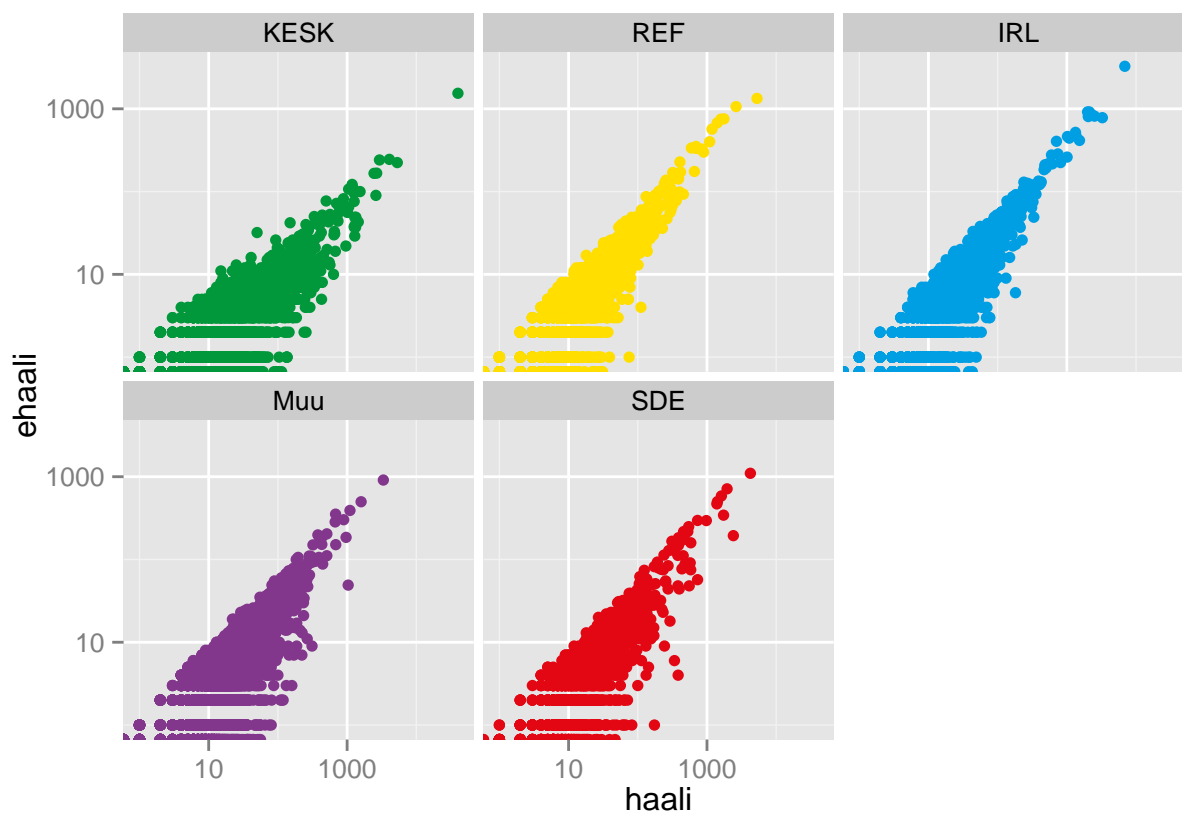


Teen plotid, mis on sarnased näidatule

```
library(dplyr)
valimised3=valimised %>%
  mutate(partei_luhend=factor(partei,
                                labels = c("KESK", "REF", "IRL", "Muu", "SDE")))

ggplot(valimised3, aes(x=haali, y=ehaali, colour=partei_luhend))+
  scale_y_log10()+
  scale_x_log10()+
  geom_point()+
  scale_colour_manual(values = c("#00983A", "#FFDE00", "#009FE3",
                                "#82368C", "#E30613"))+

  facet_wrap(~partei_luhend)+
  theme(legend.position="none")
```



```
#üritame teha nii, ert igal facetil on kõik punktid, kuid ainult
#vastava erakonna oma on värviline
ggplot(valimised3, aes(x=haali, y=ehaali, colour=partei_luhend))+
  scale_y_log10()+
  scale_x_log10()+
  geom_point(data = transform(valimised3, partei_luhend = NULL), colour="grey")+
  geom_point()+
  scale_colour_manual(values = c("#00983A", "#FFDE00", "#009FE3",
                                "#82368C", "#E30613"))+

  facet_wrap(~partei_luhend)+
  theme(legend.position="none")
```

