

# Antagningsdata

Jonas Björnerstedt

2017-11-18

```
library(tidyverse)
library(stringr)
library(knitr)
load("~/Library/Mobile Documents/com~apple~CloudDocs/Work/data/admin/ladok2.Rdata")
antagning = readRDS( "~/Library/Mobile Documents/com~apple~CloudDocs/Work/data/admin/antagning-ne.rds")
antagna = antagning %>% filter(resultat == "Antagen", antagningsår < 2017 | antagningsår == 2017 & anta
antagning18 = antagning %>% filter(antagningsår == 2018)
ladok = kursresult %>% filter(ar > 2012, is.na(progr)) %>% rename(kurs = kursid)
# ladok = kursresult %>% filter(ar > 2012, kursnamn != "Masters", is.na(progr)) %>% rename(kurs = kursid)
antagningProgram = readRDS( "~/Library/Mobile Documents/com~apple~CloudDocs/Work/data/admin/antagning.rds")
  filter(str_sub(kurs_programkod, 1,1) == "P")
platser = readRDS( "~/Library/Mobile Documents/com~apple~CloudDocs/Work/data/admin/platser.rds") %>% ar
```

## Koppling till Ladok

Hur många finns i både Ladok och NyA, samma personnummer och kurs?

```
bind_rows(
  antagna %>% semi_join(ladok, by = c("pnr")) %>% tally %>% mutate(dataset = "Nya & Ladok"),
  antagna %>% anti_join(ladok, by = c("pnr")) %>% tally %>% mutate(dataset = "Ej i Ladok"),
  ladok %>% semi_join(antagna, by = c("pnr")) %>% tally %>% mutate(dataset = "Ladok & Nya"),
  ladok %>% anti_join(antagna, by = c("pnr")) %>% tally %>% mutate(dataset = "Ej i Nya")
) %>% select(dataset,n) %>%
  kable
```

dataset	n
Nya & Ladok	1126
Ej i Ladok	484
Ladok & Nya	1216
Ej i Nya	491

Hur många finns i både Ladok och NyA, samma personnummer ?

```
bind_rows(
  antagna %>% semi_join(ladok, by = c("pnr", "kurs")) %>% tally %>% mutate(dataset = "Nya & Ladok"),
  antagna %>% anti_join(ladok, by = c("pnr", "kurs")) %>% tally %>% mutate(dataset = "Ej i Ladok"),
  ladok %>% semi_join(antagna, by = c("pnr", "kurs")) %>% tally %>% mutate(dataset = "Ladok & Nya"),
  ladok %>% anti_join(antagna, by = c("pnr", "kurs")) %>% tally %>% mutate(dataset = "Ej i Nya"),
  ladok %>% anti_join(bind_rows(antagna, antagningProgram), by = c("pnr")) %>% tally %>% mutate(datas
) %>% select(dataset,n) %>% kable
```

dataset	n
Nya & Ladok	1098
Ej i Ladok	512
Ladok & Nya	1056

dataset	n
Ej i Nya	651
Ej i Nya eller program	491

Vilka är antagna men finns ej i Ladok?

```
bara_nya = antagna %>% anti_join(ladok, by = c("pnr")) %>% mutate(antagningstermin = c("vt", "ht")[antagningstermin])
bara_nya %>% group_by(antagningsår, antagningstermin) %>% tally %>% spread(antagningstermin, n) %>% select(antagningsår, vt, ht)
```

antagningsår	vt	ht
2013	78	61
2014	31	62
2015	33	80
2016	37	44
2017	58	NA

Det finns ganska många personnummer i Ladok som inte finns i antagningsdata i Nya

```
bara_ladok = ladok %>% anti_join(antagna, by = c("pnr")) %>%
  rename(antagningsår = ar, antagningstermin = terminid) %>%
  mutate(antagningstermin = c("vt", "ht")[antagningstermin])
bara_ladok %>% group_by(antagningsår, antagningstermin) %>% tally %>%
  spread(antagningstermin, n) %>% select(antagningsår, vt, ht) %>% kable
```

antagningsår	vt	ht
2013	50	43
2014	35	36
2015	43	65
2016	31	27
2017	25	136

Dubbletter finns i antagning:

```
antagna %>% group_by(pnr, kurs) %>% tally %>% rename(x = n) %>% group_by(x) %>% tally
```

```
## # A tibble: 4 x 2
##       x     n
##   <int> <int>
## 1     1  1483
## 2     2    58
## 3     3     2
## 4     5     1
```

```
multiant = antagna %>% group_by(pnr, kurs) %>% mutate(ct = n(), la = last(ct)) %>% filter(ct > 2) %>% arrange(pnr, kurs)
```

I ladok är inga dubletter:

```
ladok %>% group_by(pnr, kurs) %>% tally %>% rename(x = n) %>% group_by(x) %>% tally
```

```
## # A tibble: 1 x 2
##       x     n
##   <int> <int>
```

## 1 1 1707