

# Analyses

*Philip & Philipp*

*14 Apr 2016*

```
if(!file.exists("data_final.rda")) {  
  source('data_combine.R')  
}  
  
# Opens data_final.rda if not already loaded as an object  
if(!exists("z.df" )) {  
  load("data_final.rda")  
}
```

## Motivation

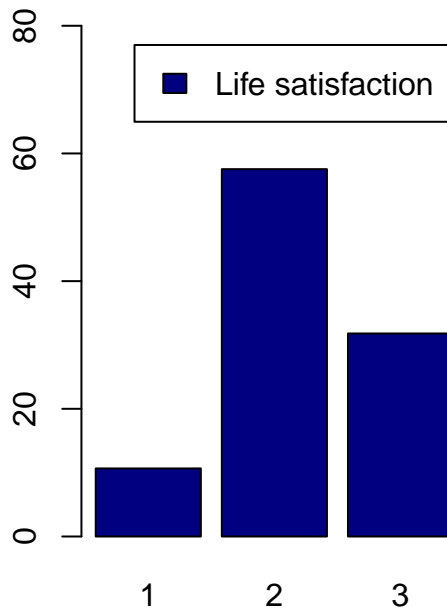
- Populate from google docs

## Descriptive results

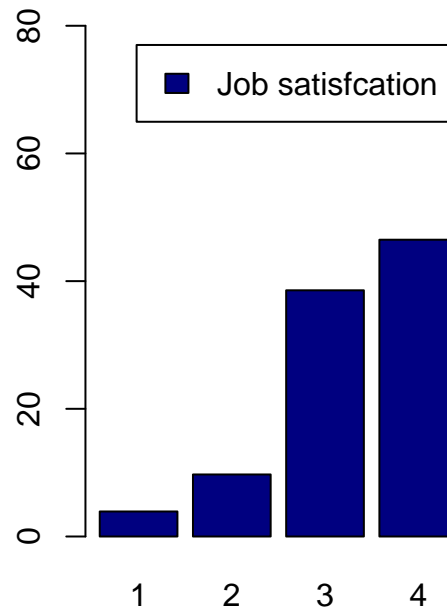
### Distribution of overall and life satisfaction

!The distribution of the well-being variables should probably be presented in a table.!

```
happy_dist <- table(z$happy)/sum(table(z$happy))*100  
satjob_dist <- table(z$satjob)/sum(table(z$happy))*100  
  
par(mfrow=c(1,2))  
  
barplot(happy_dist,  
        col=c("navyblue"),  
        xlab = "Overall life satisfaction categories",  
        legend = c("Life satisfaction"),  
        ylim = c(0, 80))  
  
barplot(satjob_dist,  
        col=c("navyblue"),  
        xlab = "Job satisfaction categories",  
        legend = c("Job satisfaction"),  
        ylim = c(0, 80))
```



Overall life satisfaction categories



Job satisfaction categories

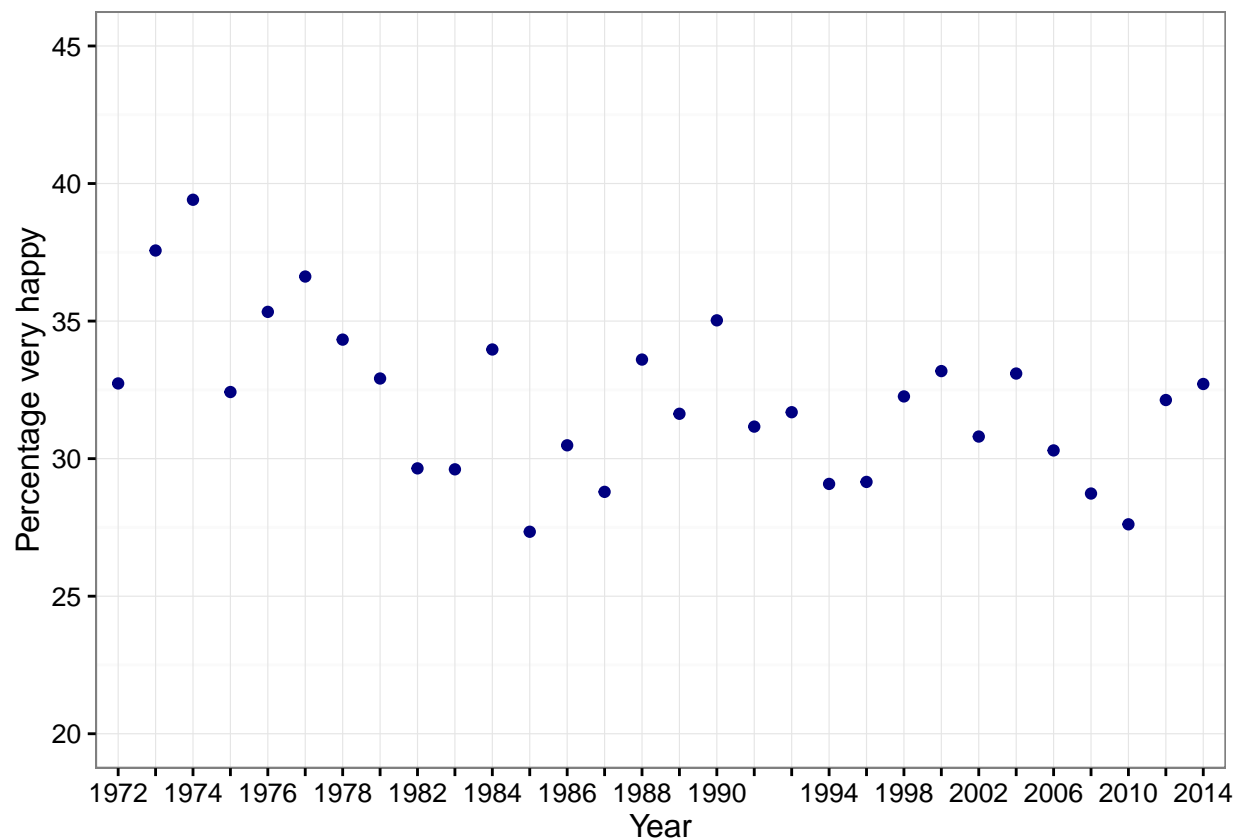
Figure x shows the distribution of survey respondents' answer to the two questions: "Taken all together, how would you say things are these days?" and "On the whole, how satisfied are you with the work you do?". The former is measured on a three point scale, of which around 60 % of the sample respond a mid-level happiness level. On the job-satisfaction variable, a majority of the sample responds either moderate or high satisfaction (score 3 and 4).

## Happiness in different survey years

The GSS is conducted between 1972 and 2014. Due to year specific events, unintended differences in how the survey was implemented or trends in overall happiness, there could be year-specific differences. Figure x shows the average share of the population who report high-happiness.

```
t <- z
t$vhappy <- t$vhappy*100
ggplot(t, aes(x=factor(year), y=vhappy)) +
  stat_summary(fun.y="mean", geom="point", col="Navyblue") +
  scale_x_discrete("Year", labels = c("1973" = "", "1975" = "", "1977" = "", "1980" = "", "1983" = "",
  scale_y_continuous("Percentage very happy") +
  expand_limits(y=c(20,45)) +
  theme_bw()
```

```
## Warning: Removed 2461 rows containing non-finite values (stat_summary).
```



The figure shows that there is considerable variation between years, and a weakly negative trend particularly between 1972 and 1983. It is not directly possible to disentangle what is caused by random noise and what is caused by structural differences, however, it signifies that it is pragmatic to control for survey year in a regression design to control for survey-year fixed effects.

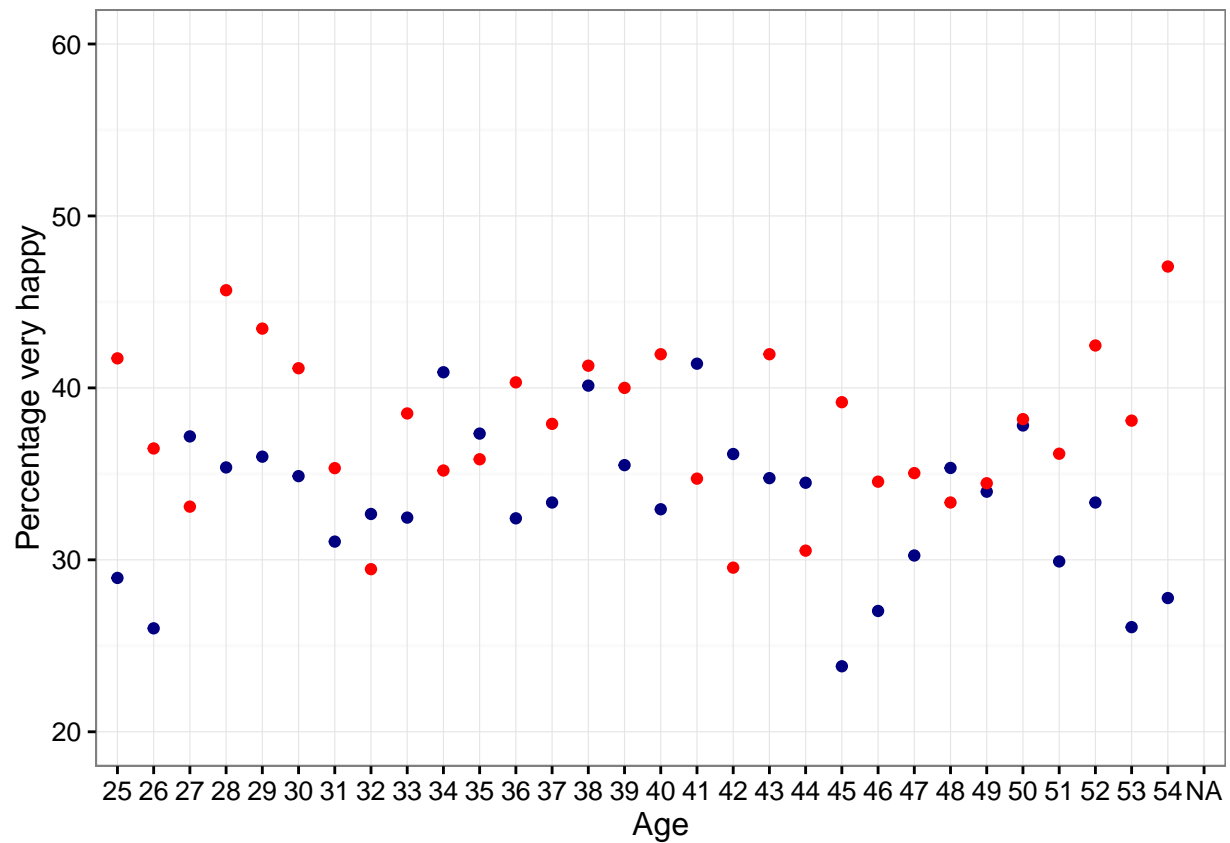
## Happiness across age for college educated men and women

Previous studies have found that there is a correlation between life satisfaction and age. However, this is not apparent. Likewise, there could be gender specific differences.

```
ggplot() +
  stat_summary(data = t[t$sex == 1 & t$educat == 4,], aes(x=factor(age), y=vhappy),
    fun.y="mean", geom="point", col="Navyblue") +
  stat_summary(data = t[t$sex == 2 & t$educat == 4,], aes(x=factor(age), y=vhappy),
    fun.y="mean", geom="point", col="Red") +
  scale_x_discrete("Age") +
  scale_y_continuous("Percentage very happy") +
  expand_limits(y=c(20,60)) +
  theme_bw()
```

```
## Warning: Removed 374 rows containing non-finite values (stat_summary).
```

```
## Warning: Removed 439 rows containing non-finite values (stat_summary).
```



College educated women has a slightly higher reported happiness overall relative to men (38% vs. 34%).

## Distribution of work-hours for full-time workers

```
hist(z$hrs1[z$working_ft==1],
     col=c("navyblue"),
     border = "White",
     xlab = "Hours worked last week (full-time employees)"
)
```

