# Visualize data

Isaac Verminck

2018-04-06

```
load(file = paste(data_path, "year_of_birth.RData", sep = "/"))
```

Visualize data

At what age do people marry?

For a start, let's not focus on any specific regions and have a look at the most recent data. How is the number of marriages distributed among all ages for all of Belgium?

```
recent_marriage_belgium <- year_of_birth %>%
  filter(year == "2014" & type == "marriage" & region == "belgium")

ggplot(data = recent_marriage_belgium,
       aes(x = age, y = quantity)) +
  geom_line(group = 1) +
  theme(axis.text.x = element_text(angle = 90, hjust = 1))
```
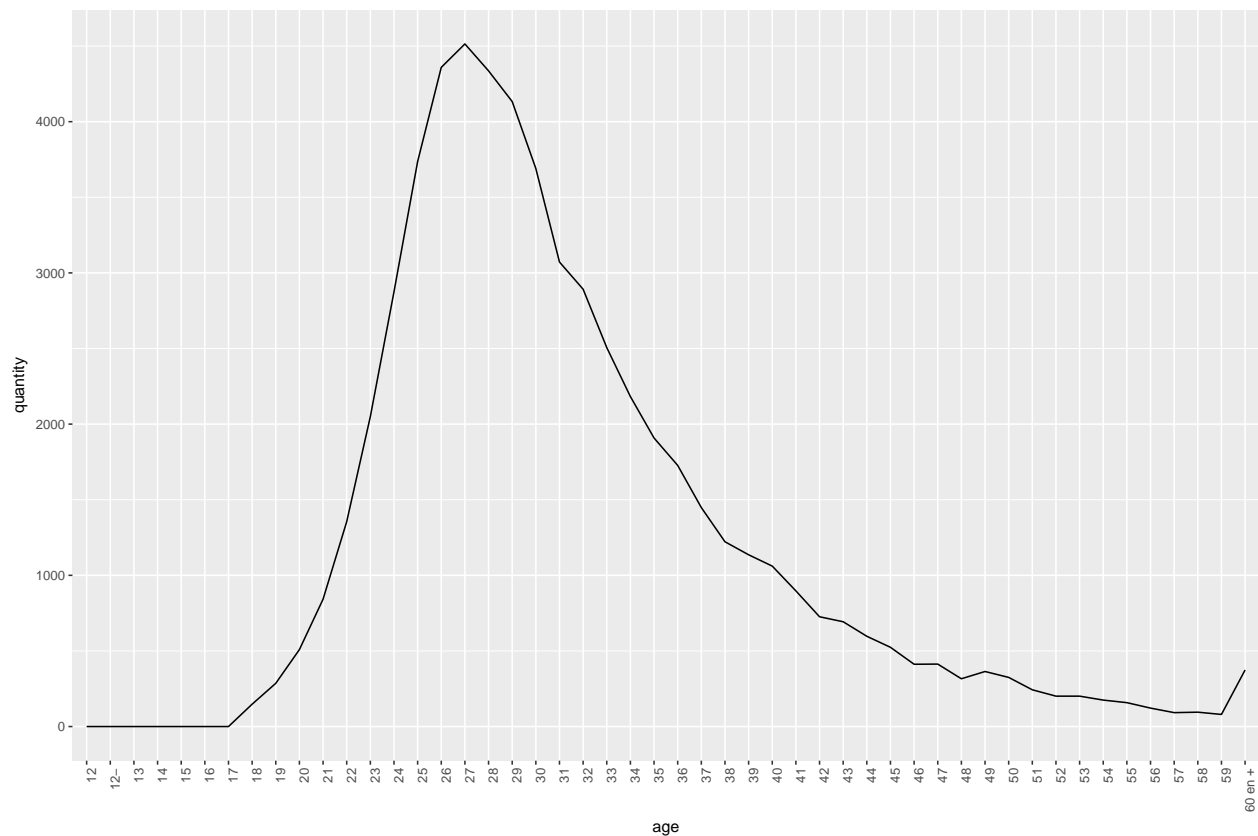

Most people marry in the end of their twenties. There's a small uptick at the end for age 60+ but that's just because multiple ages are binned there. Child marriages aren't allowed in Belgium so no marriages before the

age of 18.

Regional differences

Is there a regional difference? We remove the German community from the data set since the number of observations is too limited to make valid conclusions. Since the population of each region differs quite a lot we don't use the raw counts but instead calculate for each age group the percentage of the total number of marriages. Since we round the percentages the total of the percentages can be slightly more or less than 100% (but it's very close). We also removed the binning 60+ group and ages before 18 since there are no marriages in that age range anyway. This allows us to use age as a continuous variable.

```
# no german community and age as continuous variable
recent_marriage_regions <- year_of_birth %>%
  filter(year == "2014" & type == "marriage") %>%
  only_main_regions %>%
  age_continuous

# percentages
recent_marriage_regions <- recent_marriage_regions %>%
  percentage_of_total_by_region

# graph
ggplot(data = recent_marriage_regions,
       aes(x = age, y = percentage)) +
  geom_path(aes(color = region)) +
  theme(axis.text.x = element_text(angle = 90, hjust = 1))
```
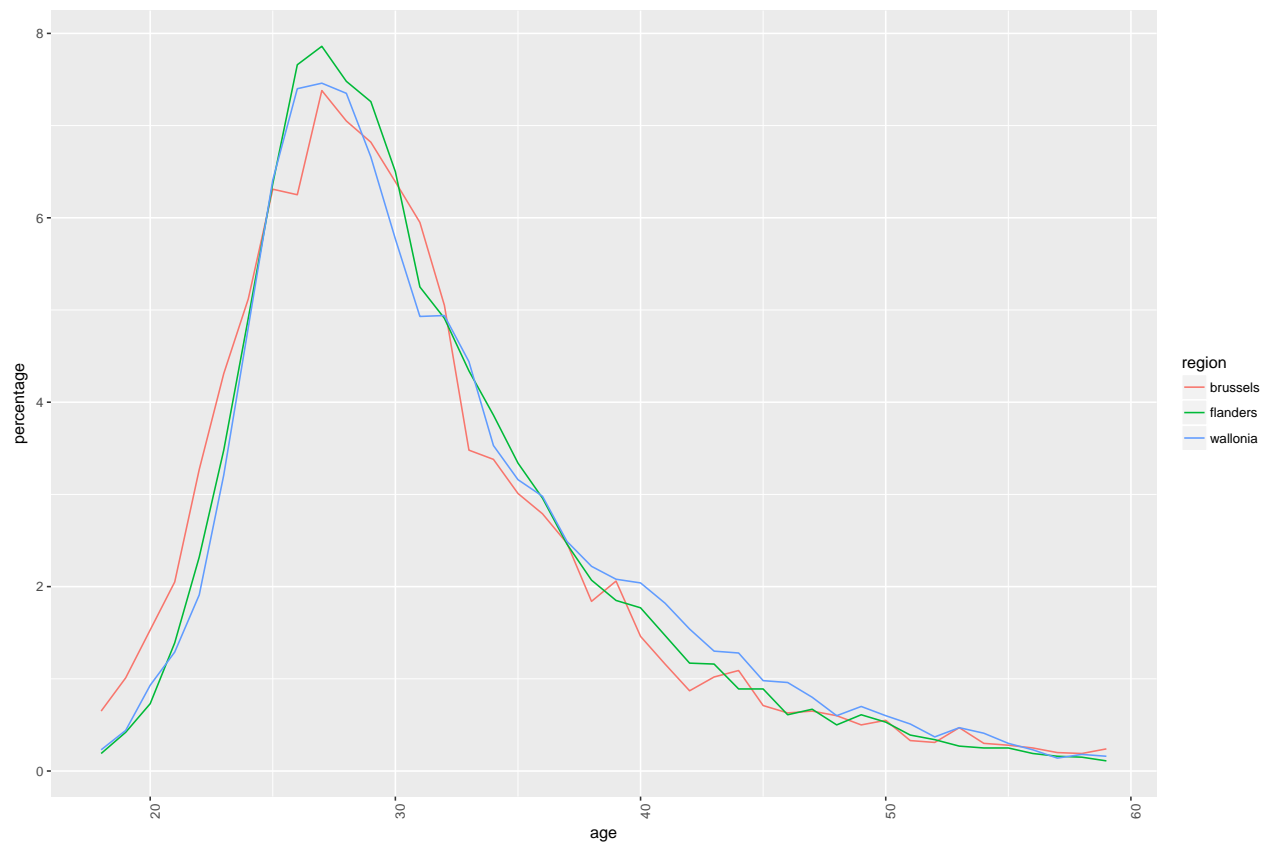
People in Brussels seem to marry at a slightly younger age then in Flanders or Wallonia. However, the difference isn't too pronounced.

Difference through time

Would there also be a change from year to year? We only have data for 2013 and 2014 so conclusions will be very limited by definition.

```r
# marriage regions
marriage_regions <- year_of_birth %>%
  filter(type == "marriage") %>%
  age_continuous %>%
  only_main_regions %>%
  percentage_of_total_by_region

# year-to-year differences
marriage_regions <- marriage_regions %>%
  year_to_year_percentage_differences

# graph
ggplot(data = marriage_regions, aes(x = age, y = difference)) +
  geom_path(aes(color = region)) +
  theme(axis.text.x = element_text(angle = 90, hjust = 1))
```
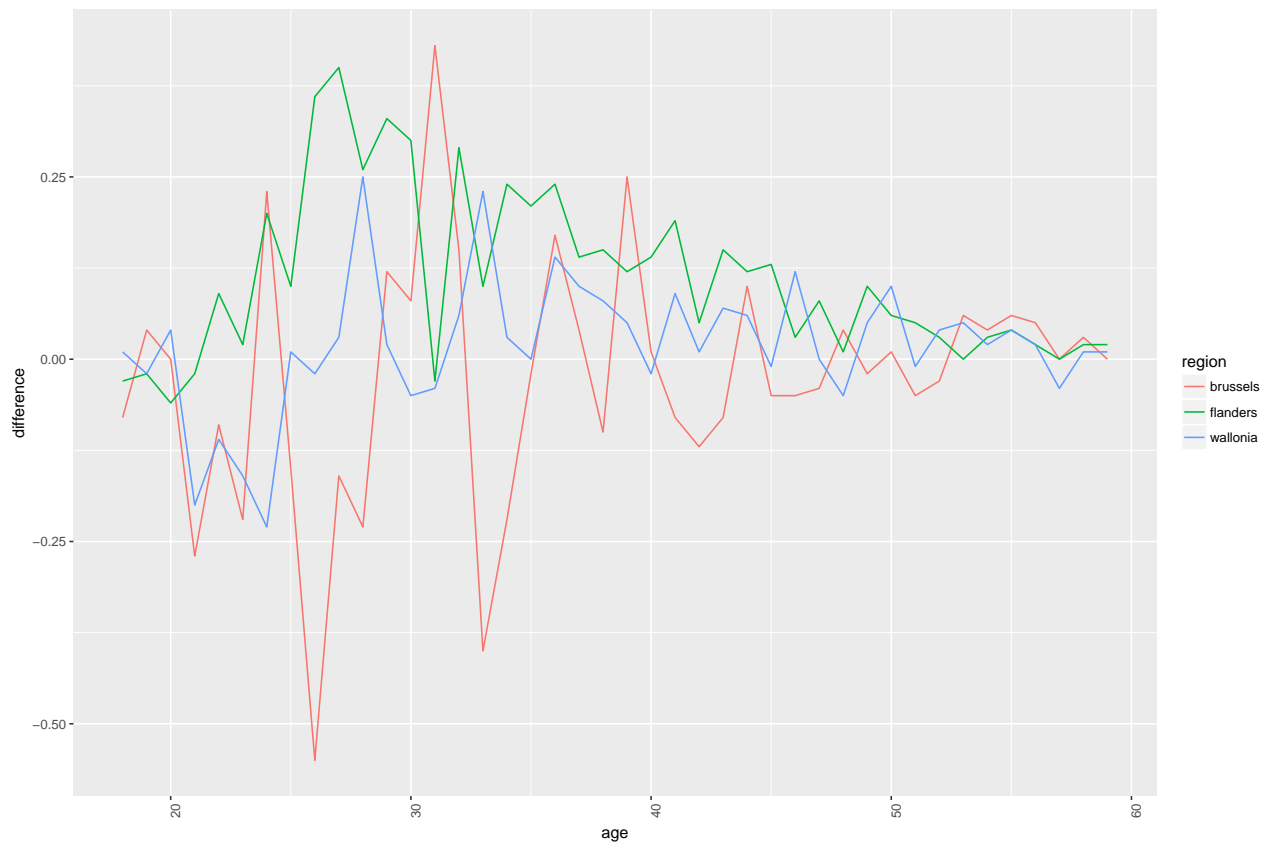

The differences from year to year seem very small and quite random so we can assume there's no difference in the age of marriage between 2013 and 2014.