Week 8 Methods: Aging

ECON 125: The Science of Population

Setup

Today, we analyze age and sex structure by country and year:

- Population counts (in 1000s) for every country-year-age-sex combination
- Every tenth year since 1950
- From United Nations World Population Prospects

Start by setting up R and loading the data

```
# Load tidyverse and clear the R environment
library(tidyverse)
rm(list=ls())

# Load dataset
country_year_age_sex <- read_csv(url("https://github.com/tomvogl/econ12

# Ask R to not use scientific notation
options(scipen = 999)
```

Glimpse

```
glimpse(country_year_age_sex)
```

```
## Rows: 382,992
## Columns: 5
## $ Country <chr> "Burundi", "Burundi", "Burundi", "Burundi", "Burundi
## $ Year    <dbl> 1950, 1950, 1950, 1950, 1950, 1950, 1950, 1950, 1950
## $ Age     <dbl> 0, 0, 1, 1, 2, 2, 3, 3, 4, 4, 5, 5, 6, 6, 7, 7, 8, 8
## $ Sex     <chr> "Males", "Females", "Males", "Females", "Males", "Fe
## $ Count   <dbl> 41.036, 41.926, 38.596, 39.626, 35.381, 36.604, 33.5
```

Setup

Set up data for plotting

Female counts to the left, male counts to the right

Simply code male counts as negative

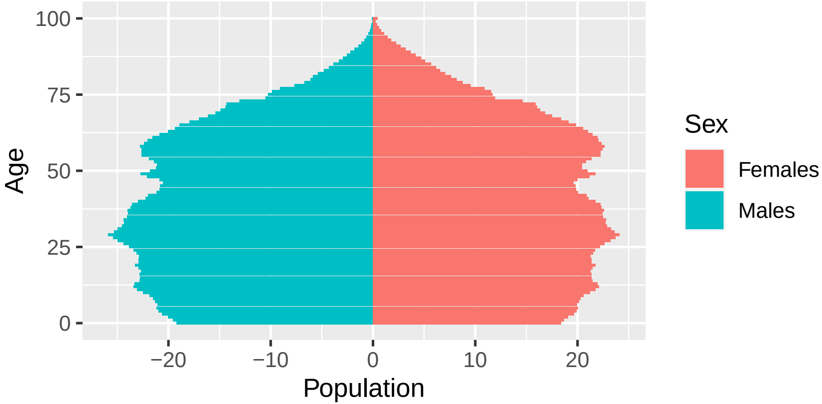
For interpretability, convert 1000s to 100,000s

```
country_year_age_sex <-
  country_year_age_sex |>
  mutate(Population = if_else(Sex=="Males", -Count/100, Count/100))
```

Population Pyramid for the US in 2020

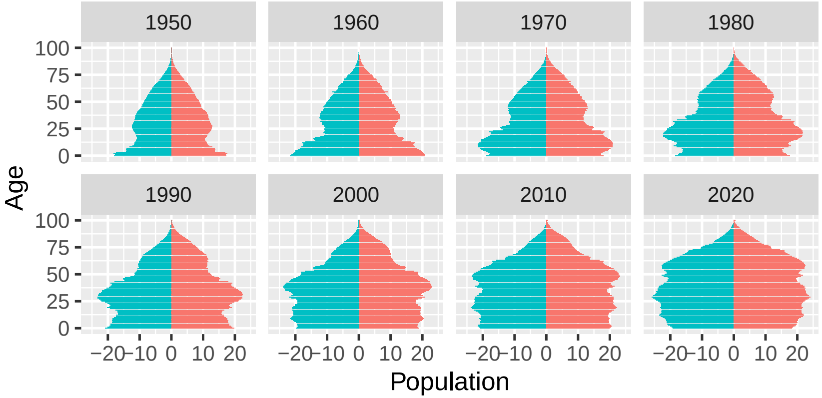
For flexibility, combine the pipe |> with ggplot()

```
country_year_age_sex |>
  filter(Country=="United States of America" & Year==2020) |>
  ggplot(aes(x = Age, y = Population, fill = Sex)) +
  geom_col() + # bar plot
  coord_flip() # horizontal bars
```



Population Pyramid for the US over Time

```
country_year_age_sex |>
  filter(Country=="United States of America") |>
  ggplot(aes(x = Age, y = Population, fill = Sex)) +
  geom_col() +
  coord_flip() +
  facet_wrap(~Year, ncol = 4, nrow = 2) + # 8 panels
  theme(legend.position = "none") # no legend
```



Population Pyramid for the US over Time

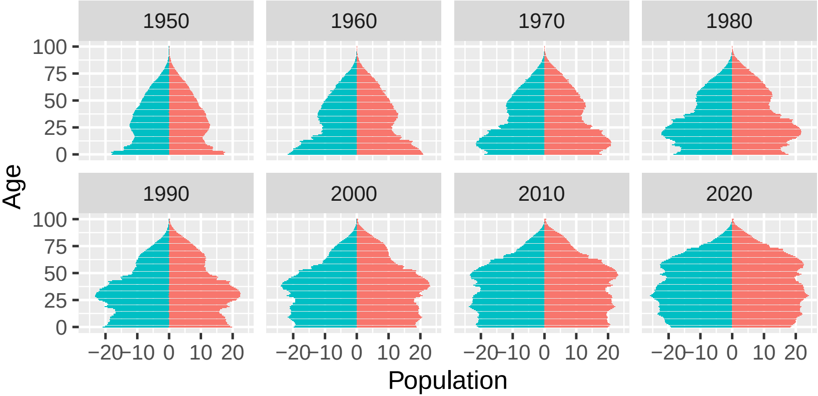
Interesting generational dynamics

Baby boom c. 1950 → many young adults c. 1980 → many children 1990

Known as **population momentum**: ↑ births because ↑ fertile women

Can also see **aging**: from rough pyramid (1950) to emerging beehive (2020)

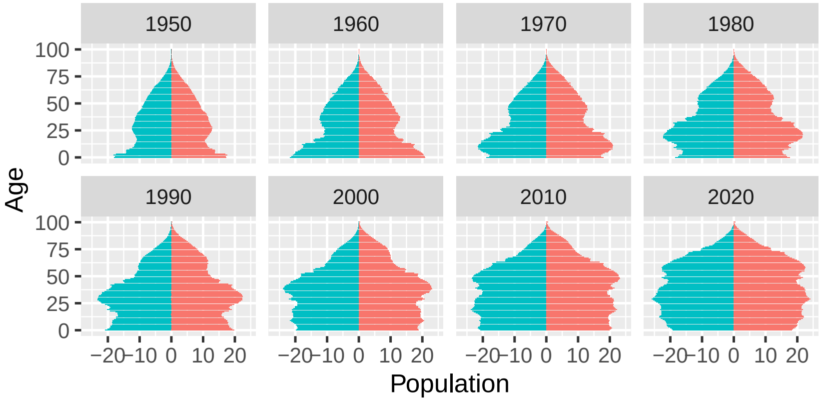
Scoops taken out of the 1950 pyramid: small cohorts born during Depression



Population Pyramid for the US over Time

If you were a marketing exec, what sorts of products would you prioritize in...

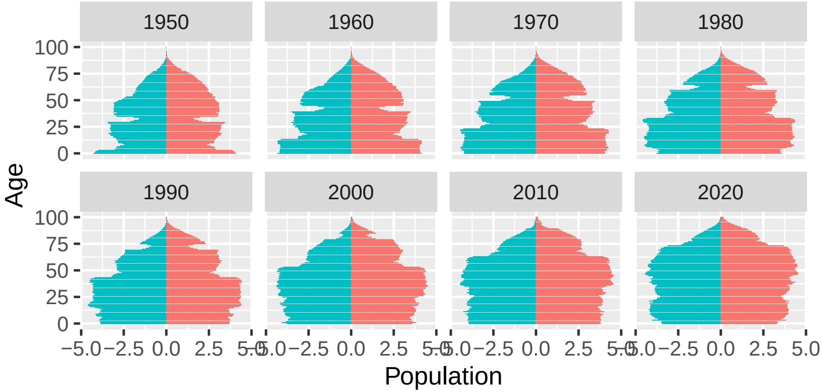
- 1950?
- 1980?
- 2010?



France

In some ways, similar to the US, but clearer effects of World Wars

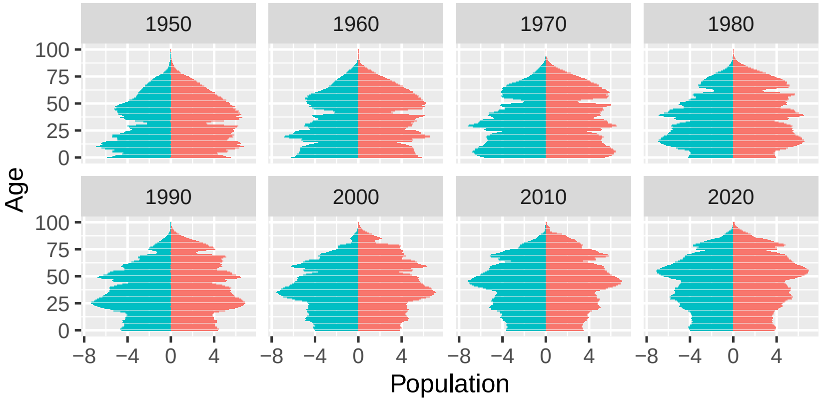
```
country_year_age_sex |>
  filter(Country=="France") |>
  ggplot(aes(x = Age, y = Population, fill = Sex)) +
  geom_col() +
  coord_flip() +
  facet_wrap(~Year, ncol = 4, nrow = 2) +
  theme(legend.position = "none")
```



Germany

Wars even more evident for Germany → sex ratio imbalances due to war deaths

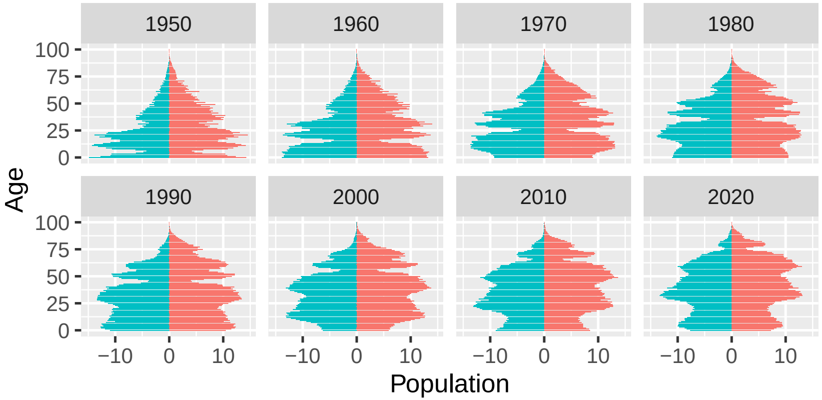
```
country_year_age_sex |>
  filter(Country=="Germany") |>
  ggplot(aes(x = Age, y = Population, fill = Sex)) +
  geom_col() +
  coord_flip() +
  facet_wrap(~Year, ncol = 4, nrow = 2) +
  theme(legend.position = "none")
```



Russia

Russia's case especially messy

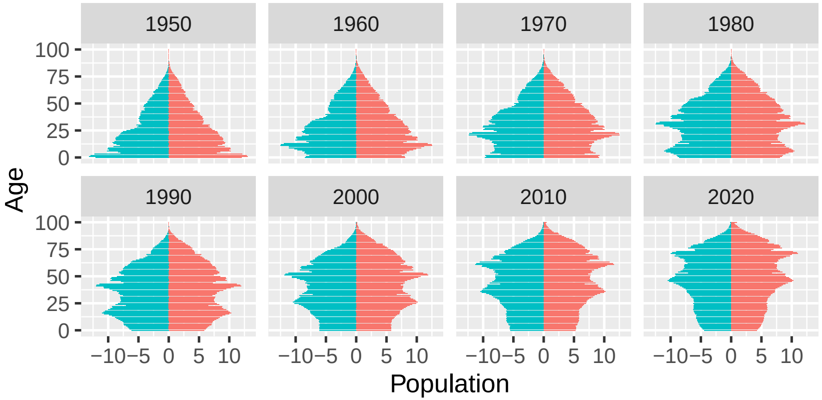
```
country_year_age_sex |>
  filter(Country=="Russian Federation") |>
  ggplot(aes(x = Age, y = Population, fill = Sex)) +
  geom_col() +
  coord_flip() +
  facet_wrap(~Year, ncol = 4, nrow = 2) +
  theme(legend.position = "none")
```



Japan

Can see WWII in Japan too; also look for Year of the Fire Horse (1966)

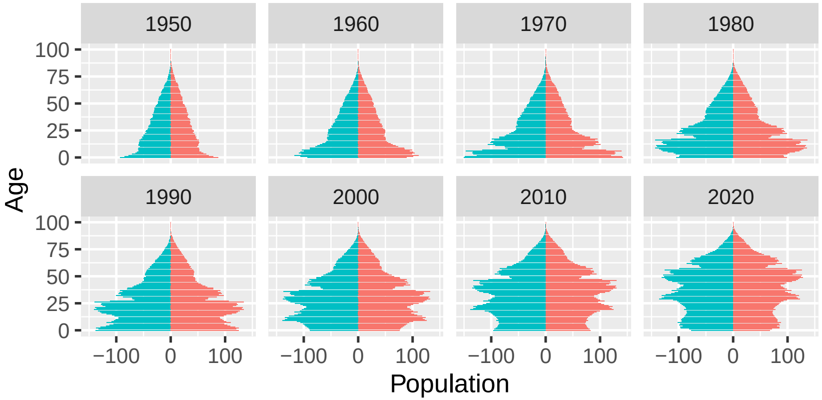
```
country_year_age_sex |>
  filter(Country=="Japan") |>
  ggplot(aes(x = Age, y = Population, fill = Sex)) +
  geom_col() +
  coord_flip() +
  facet_wrap(~Year, ncol = 4, nrow = 2) +
  theme(legend.position = "none")
```



China

Can see Great Leap Forward (c. 1960), and then fertility decline in 1970s

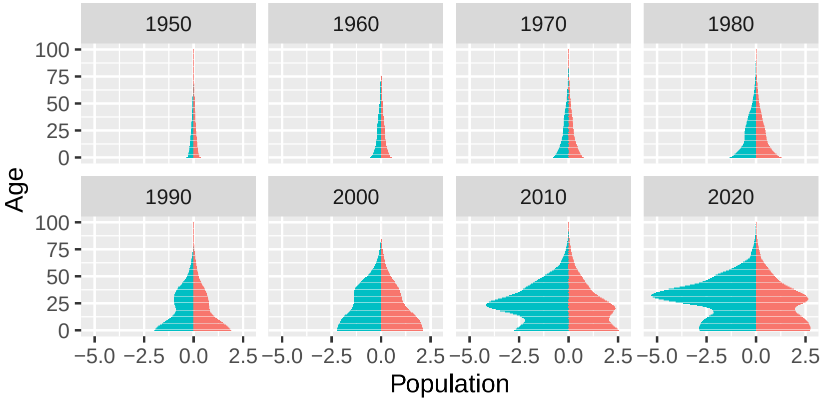
```r
country_year_age_sex |>
  filter(Country=="China") |>
  ggplot(aes(x = Age, y = Population, fill = Sex)) +
  geom_col() +
  coord_flip() +
  facet_wrap(~Year, ncol = 4, nrow = 2) +
  theme(legend.position = "none")
```



Saudi Arabia

Lots of male migrant workers

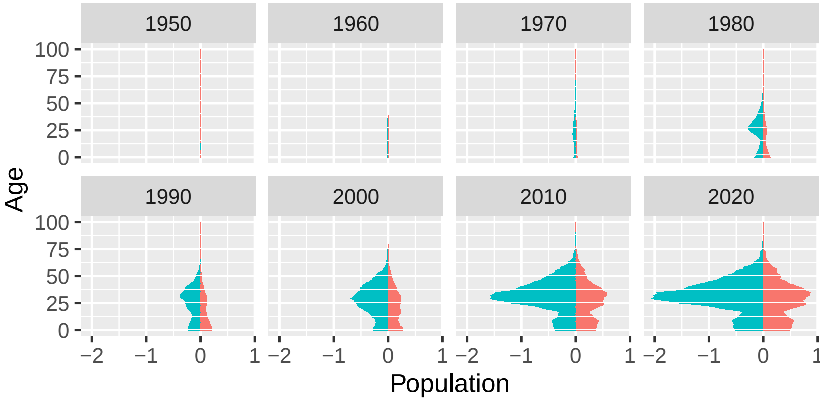
```
country_year_age_sex |>
  filter(Country=="Saudi Arabia") |>
  ggplot(aes(x = Age, y = Population, fill = Sex)) +
  geom_col() +
  coord_flip() +
  facet_wrap(~Year, ncol = 4, nrow = 2) +
  theme(legend.position = "none")
```



United Arab Emirates

Even more extreme than Saudi

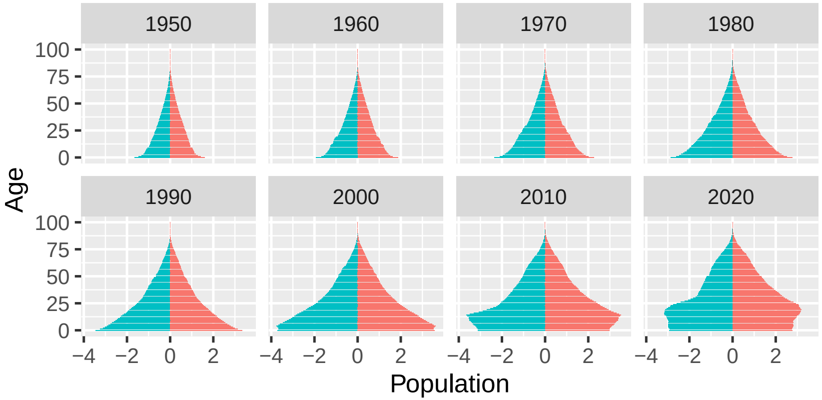
```
country_year_age_sex |>
  filter(Country=="United Arab Emirates") |>
  ggplot(aes(x = Age, y = Population, fill = Sex)) +
  geom_col() +
  coord_flip() +
  facet_wrap(~Year, ncol = 4, nrow = 2) +
  theme(legend.position = "none")
```



Nepal

Nepal today sends many male migrant workers

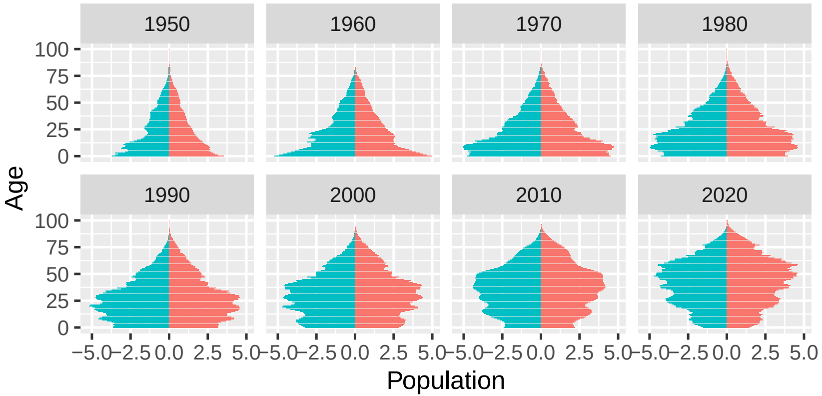
```
country_year_age_sex |>
  filter(Country=="Nepal") |>
  ggplot(aes(x = Age, y = Population, fill = Sex)) +
  geom_col() +
  coord_flip() +
  facet_wrap(~Year, ncol = 4, nrow = 2) +
  theme(legend.position = "none")
```



South Korea

South Korea gets us back to talking about aging - from pyramid to beehive

```
country_year_age_sex |>
  filter(Country=="Republic of Korea") |>
  ggplot(aes(x = Age, y = Population, fill = Sex)) +
  geom_col() +
  coord_flip() +
  facet_wrap(~Year, ncol = 4, nrow = 2) +
  theme(legend.position = "none")
```



Dependency ratios

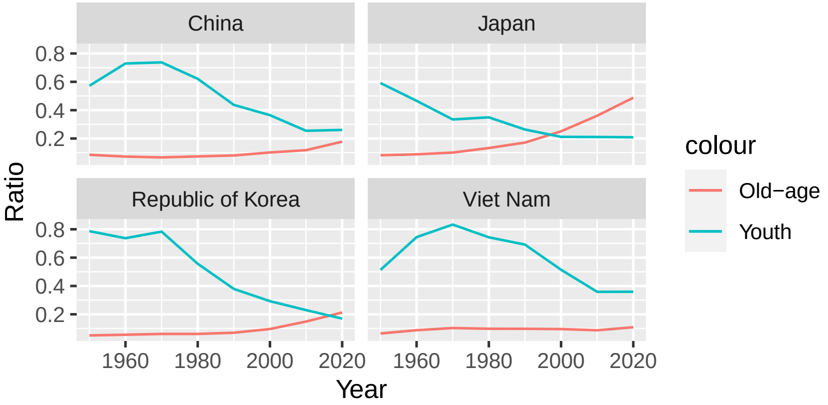
We can summarize the pyramids with dependency ratios (youth and old-age)

Necessitates switch to country-year data

```
country_year <-
  country_year_age_sex |>
  mutate(P_0_14 = if_else(Age<15, Count, 0),
         P_15_64 = if_else(Age>=15&Age<65, Count, 0),
         P_65plus = if_else(Age>=65, Count, 0)) |>
  group_by(Country, Year) |>
  summarise(Pop_0_14 = sum(P_0_14),
            Pop_15_64 = sum(P_15_64),
            Pop_65plus = sum(P_65plus)) |>
  mutate(OldRatio = Pop_65plus/Pop_15_64,
         YoungRatio = Pop_0_14/Pop_15_64,
         DepRatio = OldRatio + YoungRatio)
```

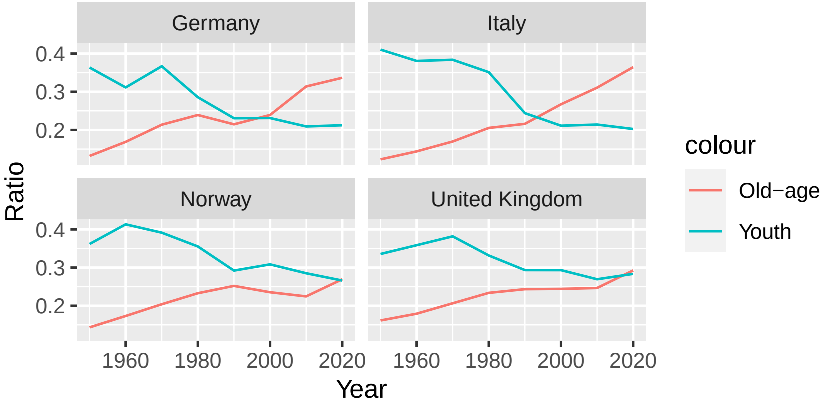
Dependency Ratios in 4 Asian Countries

```r
country_year |>
  filter(Country=="China"|Country=="Japan"|
         Country=="Republic of Korea"|Country=="Viet Nam") |>
  ggplot(aes(x = Year)) +
  geom_line(aes(y = OldRatio, color = "Old-age")) +
  geom_line(aes(y = YoungRatio, color = "Youth")) +
  labs(y = "Ratio", x = "Year") +
  facet_wrap(~Country)
```



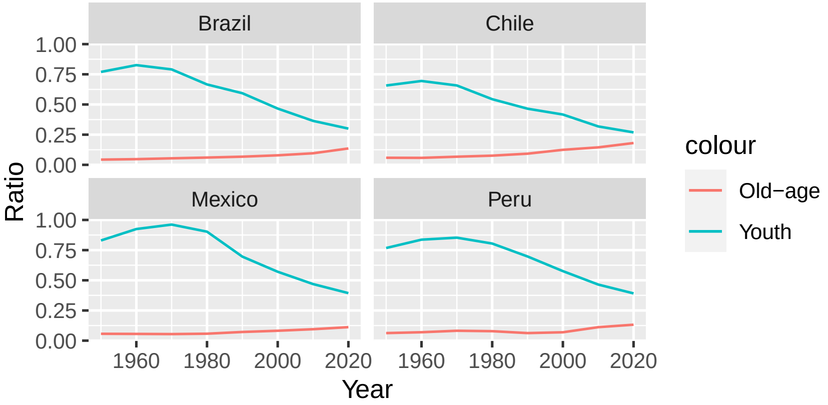
Dependency Ratios in 4 European Countries

```r
country_year |>
  filter(Country=="Germany"|Country=="Italy"|
         Country=="Norway"|Country=="United Kingdom") |>
  ggplot(aes(x = Year)) +
  geom_line(aes(y = OldRatio, color = "Old-age")) +
  geom_line(aes(y = YoungRatio, color = "Youth")) +
  labs(y = "Ratio", x = "Year") +
  facet_wrap(~Country)
```



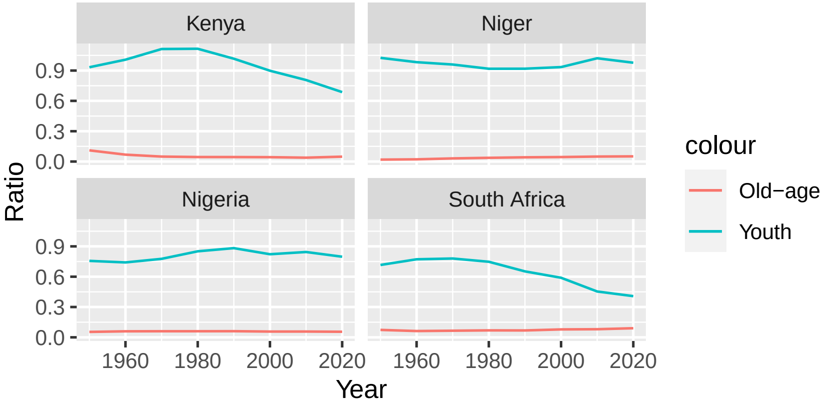
Dependency Ratios in 4 Latin American Countries

```r
country_year |>
  filter(Country=="Brazil"|Country=="Chile"|
         Country=="Mexico"|Country=="Peru") |>
  ggplot(aes(x = Year)) +
  geom_line(aes(y = OldRatio, color = "Old-age")) +
  geom_line(aes(y = YoungRatio, color = "Youth")) +
  labs(y = "Ratio", x = "Year") +
  facet_wrap(~Country)
```



Dependency Ratios in 4 African Countries

```r
country_year |>
  filter(Country=="Kenya"|Country=="Niger"|
         Country=="Nigeria"|Country=="South Africa") |>
  ggplot(aes(x = Year)) +
  geom_line(aes(y = OldRatio, color = "Old-age")) +
  geom_line(aes(y = YoungRatio, color = "Youth")) +
  labs(y = "Ratio", x = "Year") +
  facet_wrap(~Country)
```



Adding Fertility

Let's add fertility to the country-year data frame

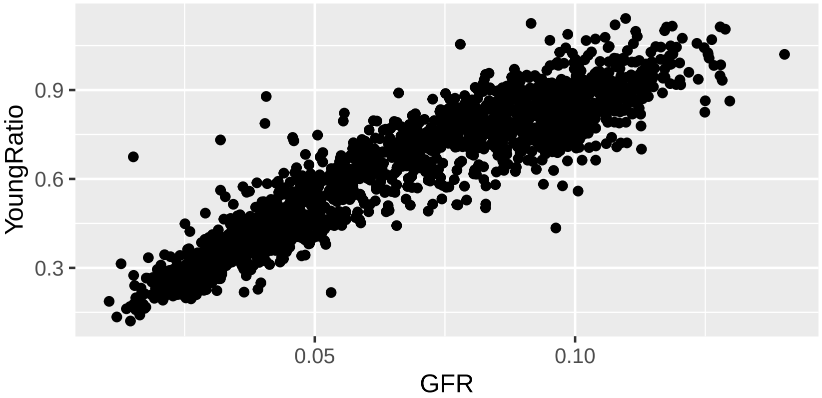
Approximate general fertility rate as $GFR = \frac{Infants}{Women15-44}$

(Why is this an approximation?)

```
country_year <-
  country_year_age_sex |>
  mutate(Infants = if_else(Age==0, Count, 0),
         Women_15_44 = if_else(Age>=15&Age<45, Count, 0)) |>
  group_by(Country, Year) |>
  summarise(GFR = sum(Infants)/sum(Women_15_44)) |>
  left_join(country_year, by=c("Country", "Year"))
```

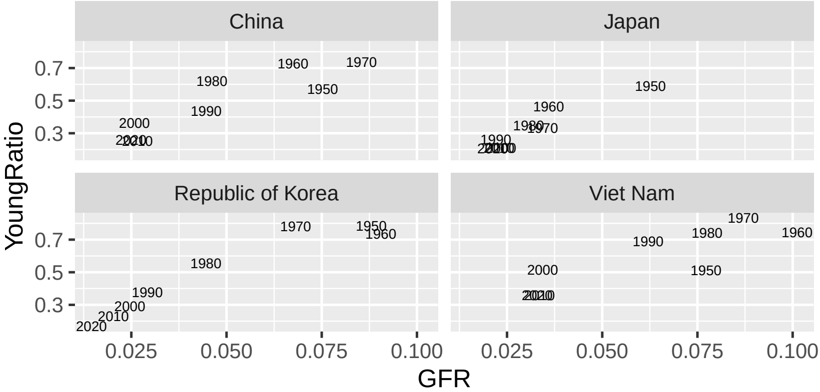
Relationship Between Fertility and Youth Dependency Ratio

```r
ggplot(country_year, aes(x = GFR, y=YoungRatio)) +
  geom_point()
```



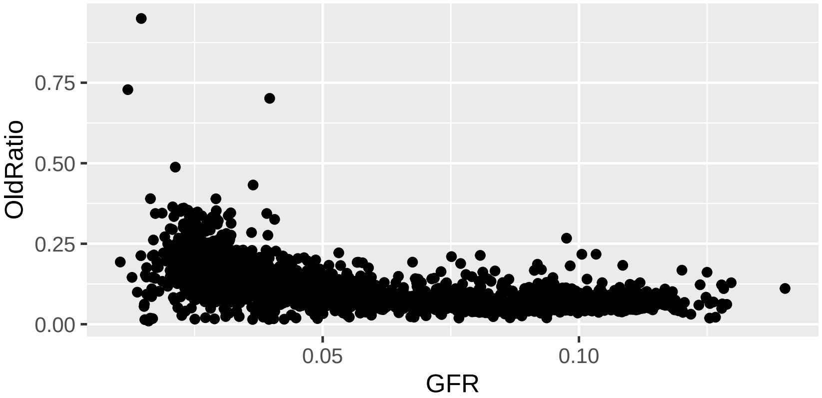
Relationship over Time in 4 Asian Countries

```
country_year |>
  filter(Country=="China"|Country=="Japan"|
         Country=="Republic of Korea"|Country=="Viet Nam") |>
  ggplot(aes(x = GFR, y=YoungRatio, label = Year)) +
  geom_text(size = 2) +
  facet_wrap(~Country)
```



Relationship Between Fertility and Old-Age Dependency Ratio

```r
ggplot(country_year, aes(x = GFR, y=OldRatio)) +
  geom_point()
```



Relationship Between Fertility and Overall Dependency Ratio

```r
ggplot(country_year, aes(x = GFR, y=DepRatio)) +
  geom_point()
```