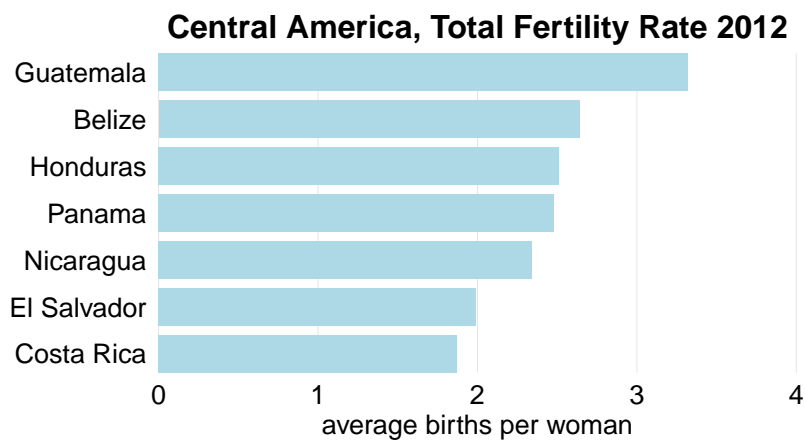


ggplot2 code for select graphs

Bar plots	1
Box plots	3
Scatterplots	4
Color scatterplots	5
Parallel coordinate plots	7
Grouped bar charts	9
Divided bar charts	10
Controlling text elements with theme() in ggplot2	11

Bar plots

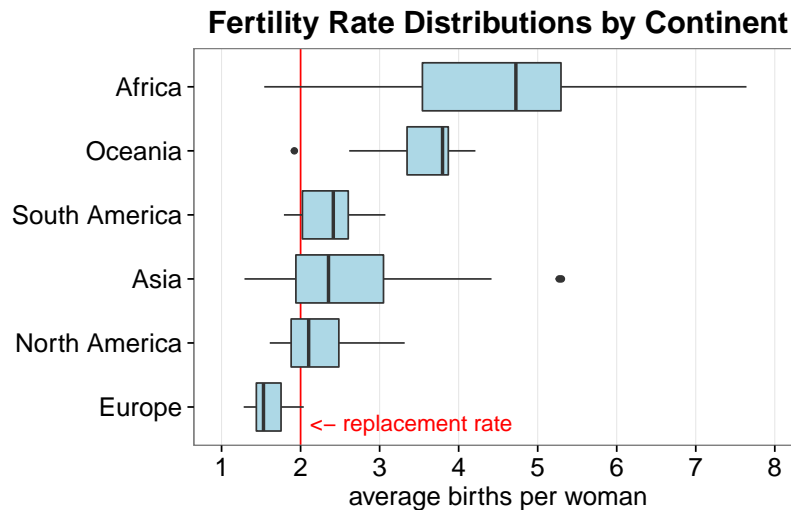


```

#+ fig.width = 7, fig.height = 5
library(ggplot2)
central <- factor(c("Costa Rica", "El Salvador",
                    "Nicaragua", "Panama", "Honduras",
                    "Belize", "Guatemala"))
centraltfr <- c(1.87, 1.99, 2.34, 2.48, 2.51, 2.64, 3.32)
data <- data.frame(COUNTRY = central, TFR = centraltfr)
g <- ggplot(data, aes(COUNTRY, TFR))
g + geom_bar(stat="identity", width = .8,
             fill = "lightblue") +
  coord_flip(expand = FALSE) +
  scale_x_discrete(limits = data$COUNTRY) +
  scale_y_continuous(limits = c(0, 4)) +
  theme_bw(16) +
  theme(aspect.ratio = .5,
        axis.line = element_blank(),
        axis.text = element_text(size = rel(1)),
        axis.ticks = element_blank(),
        panel.border = element_blank(),
        panel.grid.major = element_line(colour = "grey90"),
        panel.grid.major.y = element_blank(),
        panel.grid.minor = element_blank(),
        plot.title = element_text(face = "bold")) +
  ggtitle("Central America, Total Fertility Rate 2012") +
  ylab("average births per woman") + xlab("")

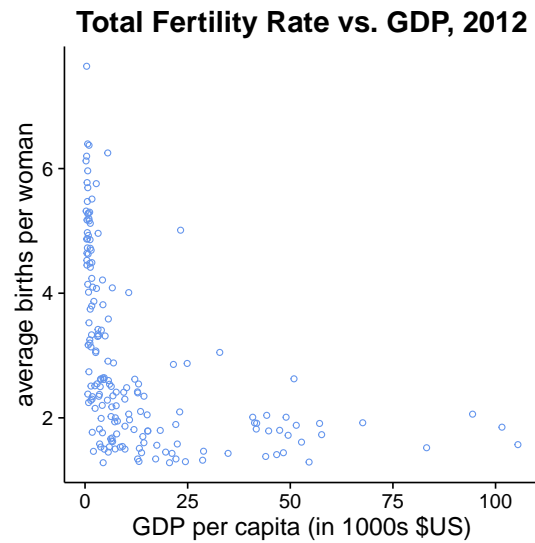
```

Box plots



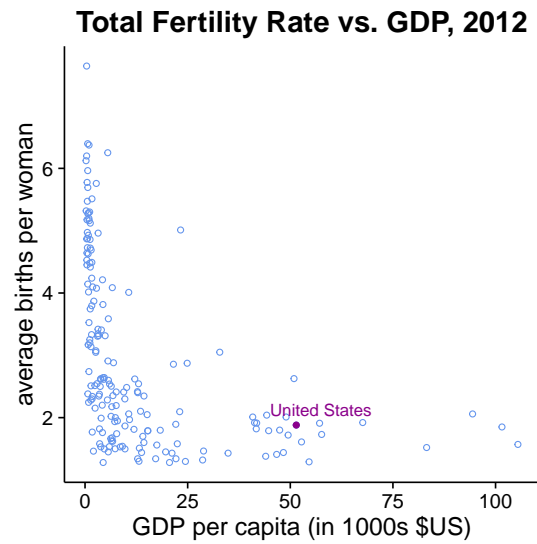
```
#+ fig.width = 7, fig.height = 4.5
library(ggplot2)
data <- read.csv("countries2012.csv")
g <- ggplot(data, aes(x = reorder(CONTINENT, TFR, median),
                       y = TFR))
g + geom_hline(yintercept = 2, color = "red") +
  geom_boxplot(fill="lightblue") +
  annotate("text", x = .75, y = 3.4,
           label = "<- replacement rate", color = "red",
           size = 5) +
  coord_flip() +
  scale_y_continuous(limits = c(1,8), breaks = 1:8) +
  theme_bw(16) +
  theme(axis.text = element_text(size = rel(1)),
        panel.grid.major.y = element_blank(),
        panel.grid.minor = element_blank(),
        plot.title = element_text(face = "bold")) +
  ggtitle("Fertility Rate Distributions by Continent") +
  xlab(NULL) + ylab("average births per woman")
```

Scatterplots

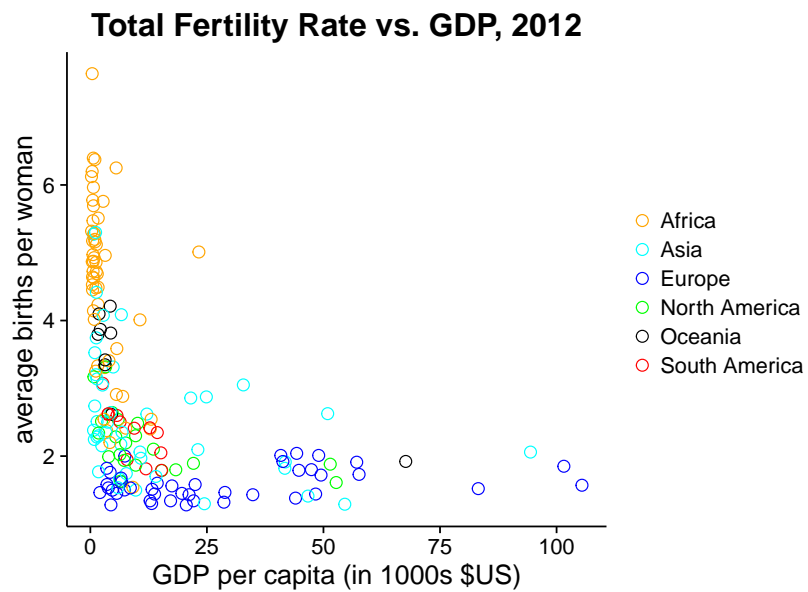


```
#+ fig.width = 5, fig.height = 5
library(ggplot2)
data <- read.csv("countries2012.csv")
g <- ggplot(data, aes(x = GDP/1000, y = TFR))
g + geom_point(shape = 1, color = "cornflowerblue") +
  theme_classic(16) +
  theme(plot.title = element_text(face = "bold")) +
  ggtitle("Total Fertility Rate vs. GDP, 2012") +
  xlab("GDP per capita (in 1000s $US)") +
  ylab("average births per woman")
```

Scatterplot with one label



Color scatterplots

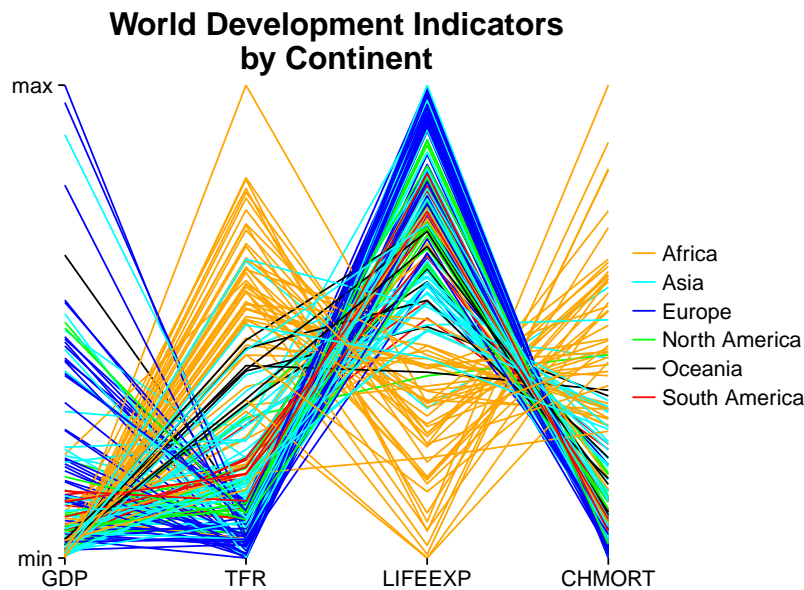


```

#+ fig.width = 7, fig.height = 5
library(ggplot2)
data <- read.csv ("countries2012.csv")
colors6 <- c("orange","cyan","blue","green",
             "black", "red")
g <- ggplot(data, aes(GDP/1000, TFR, color = CONTINENT))
g + geom_point(shape = 1, size = 3) +
  scale_color_manual(values = colors6) +
  theme_classic(16) +
  theme(plot.title = element_text(face = "bold"),
        legend.title = element_blank()) +
  ggtitle("Total Fertility Rate vs. GDP, 2012") +
  xlab ("GDP per capita (in 1000s $US)") +
  ylab ("average births per woman")

```

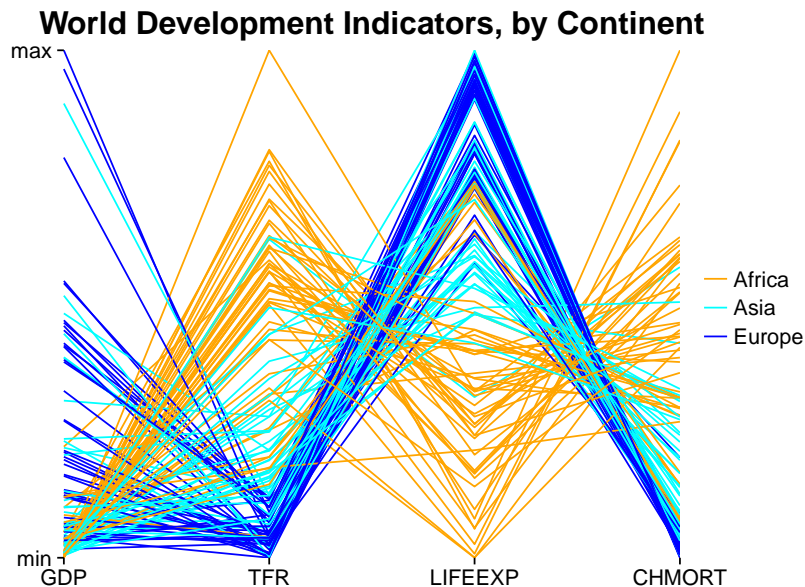
Parallel coordinate plots



```
#+ fig.width = 7, fig.height = 5
library(ggplot2)
library(tidyr)
rescale <- function(x) {(x-min(x))/(max(x)-min(x))}
data <- read.csv("countries2012.csv")
data[,3:6] <- lapply(data[,3:6], rescale)
x <- gather(data, key, value, -COUNTRY, -CONTINENT)
colors6 <- c("orange", "cyan", "blue", "green",
             "black", "red")
g <- ggplot(x, aes(x = key, y = value, group = COUNTRY,
                  color = CONTINENT))
g + geom_line() + coord_cartesian(expand = FALSE) +
  scale_y_continuous(breaks = 0:1,
                    labels = c("min", "max")) +
  scale_color_manual(values = colors6) +
  theme_classic(16) +
  theme(axis.line = element_blank(),
        legend.title = element_blank(),
        plot.title = element_text(face = "bold")) +
  ggtitle("World Development Indicators\nby Continent") +
```

```
xlab(NULL) + ylab(NULL)
```

Three continents only



```
#+ fig.width = 7, fig.height = 5
library(ggplot2)
library(tidyr)
library(dplyr)
rescale <- function(x) {(x-min(x))/(max(x)-min(x))}
data <- read.csv("countries2012.csv")
data[,3:6] <- lapply(data[,3:6], rescale)
x <- gather(data, key, value, -COUNTRY, -CONTINENT)
colors6 <- c("orange", "cyan", "blue", "green",
             "black", "red")
colors3 <- colors6[1:3]
x <- x %>% filter(CONTINENT %in% c("Africa", "Asia",
                                "Europe"))
g <- ggplot(x, aes(x = key, y = value, group = COUNTRY,
                  color = CONTINENT))
g + geom_line() + coord_cartesian(expand = FALSE) +
  scale_y_continuous(breaks = 0:1,
```

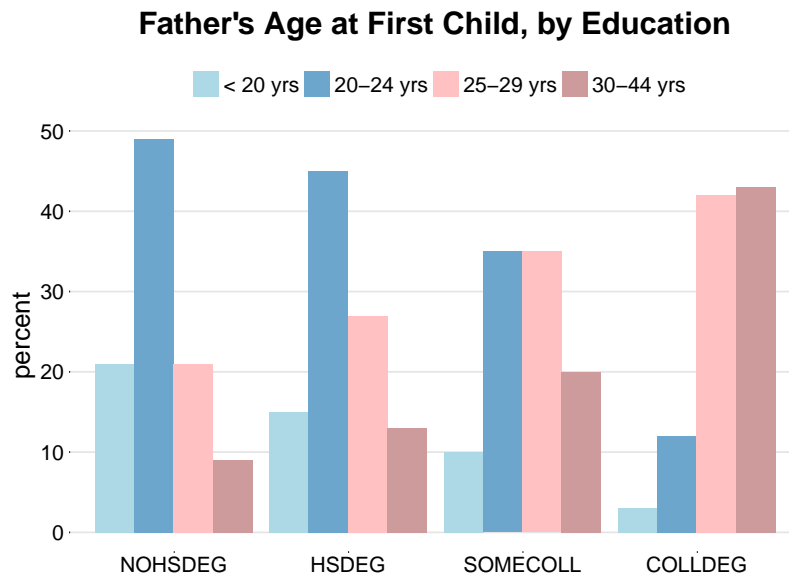


```

      labels = c("min", "max")) +
scale_color_manual(values = colors3) +
theme_classic(16) +
theme (axis.line = element_blank(),
      legend.title = element_blank(),
      plot.title = element_text(face = "bold")) +
ggtitle("World Development Indicators, by Continent") +
xlab(NULL) + ylab(NULL)

```

Grouped bar charts



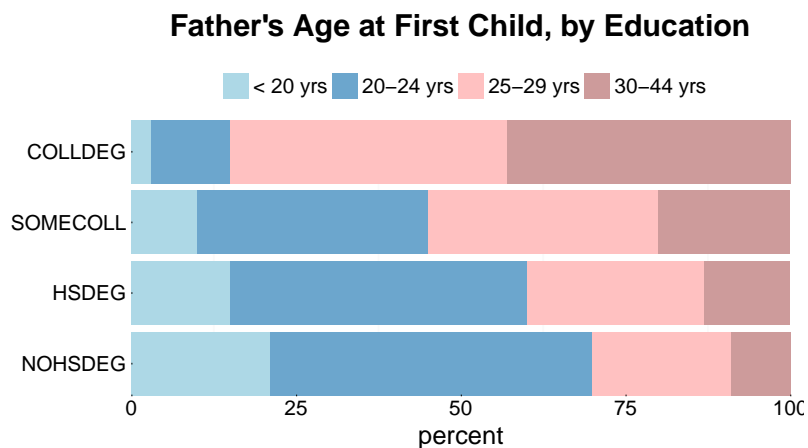
```

#+ fig.width = 7, fig.height = 5
library(tidyr)
library(ggplot2)
fathers <- read.table("fathers.txt")
data <- gather(fathers, key = EDUCATION, value = PERCENT,
               -AGE)
colors4 <- c("lightblue", "skyblue3", "rosybrown1",
             "rosybrown3")
g <- ggplot(data, aes(x = EDUCATION, y = PERCENT,
                     fill = AGE))

```

```
g + geom_bar(stat = "identity", position = "dodge") +
  scale_fill_manual(values = colors4) + theme_classic(16) +
  theme(axis.line = element_blank(),
        axis.ticks.length = unit(0, "cm"),
        panel.grid.major = element_line(colour = "grey90"),
        panel.grid.major.x = element_blank(),
        legend.position = "top",
        legend.title = element_blank(),
        plot.title = element_text(face = "bold")) +
  ggtitle("Father's Age at First Child, by Education") +
  xlab(NULL) + ylab("percent")
```

Divided bar charts



```
#+ fig.width = 7, fig.height = 4
library(ggplot2)
library(tidyr)
fathers <- read.table("fathers.txt")
data <- gather(fathers, key = EDUCATION, value = PERCENT,
               -AGE)
colors4 <- c("lightblue", "skyblue3", "rosybrown1",
             "rosybrown3")
g <- ggplot(data, aes(x = EDUCATION, y = PERCENT,
                     fill = AGE))
```

```

g + geom_bar(stat = "identity") +
  coord_flip(expand = FALSE) +
  scale_fill_manual(values = colors4) + theme_bw(16) +
  theme(panel.grid.major = element_blank(),
        panel.border = element_blank(),
        legend.position="top",
        legend.title = element_blank(),
        legend.key = element_blank(),
        axis.ticks.length = unit(0, "cm"),
        plot.title = element_text(face = "bold")) +
  ggtitle("Father's Age at First Child, by Education") +
  xlab(NULL) + ylab("percent")

```

Controlling text elements with theme() in ggplot2

Font family, face, size, and color of text can be set with:

theme(... = element_text (...)). See previous examples.

