UPPSALA
UNIVERSITET

Uppsala Child & Baby Lab

Common plots for Psychology using `ggplot2`

R for Psychology Research

Marcus Lindskog, Docent
2019-10-09

Overview

1. Graph for t-tests.
2. Graph for correlations.
3. Graph for ANOVA-like designs.
4. Graph for multiple regressions.

Graphs for t-tests

Single sample t-tests

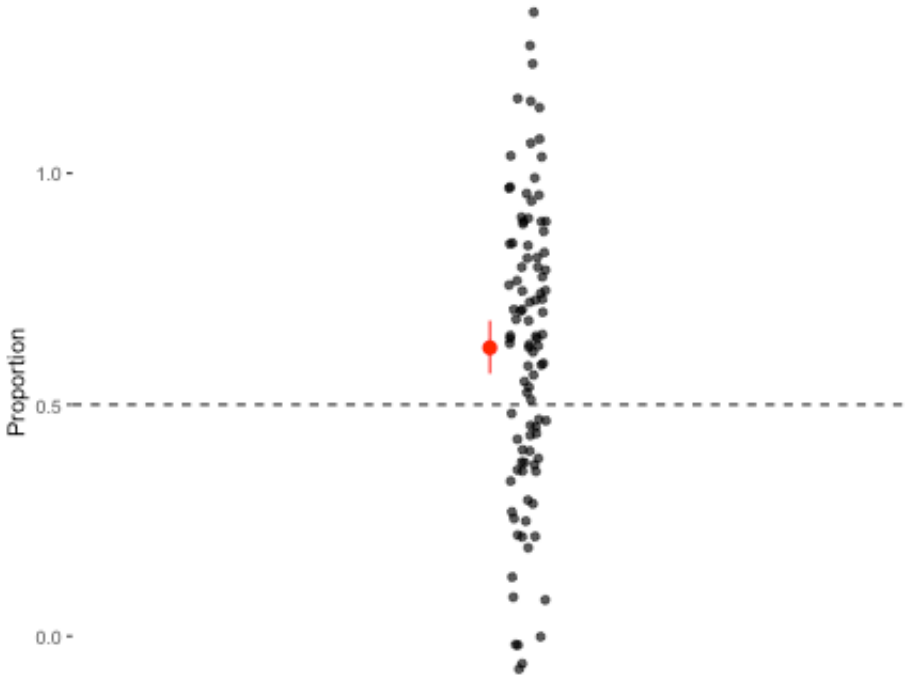
We need some data

```
# code chunk here
t_test_data <- rnorm(100, .6, .3)
data_df <- tibble(dat = t_test_data)
```

Test the data against .5

```
ggplot(data = data_df, aes(x = 1, y = dat)) +
  geom_jitter(aes(x= 1.02), width = .01,
              alpha = .7) +
  stat_summary(fun.data = "mean_cl_boot",
               size = .5, colour = "red") +
  ylim(c(-.3,1.4))+
  xlim(c(0.8,1.2))+
  geom_hline(yintercept = .5, linetype =
"dashed") +
  theme_bw() +
  labs(y = "Proportion", x = "") +
  theme(axis.text.x = element_blank(),
        panel.grid = element_blank(),
        panel.border = element_blank(),
        axis.ticks.x = element_blank())
```

Test the data against .5



What did that code do?

Two group t-tests

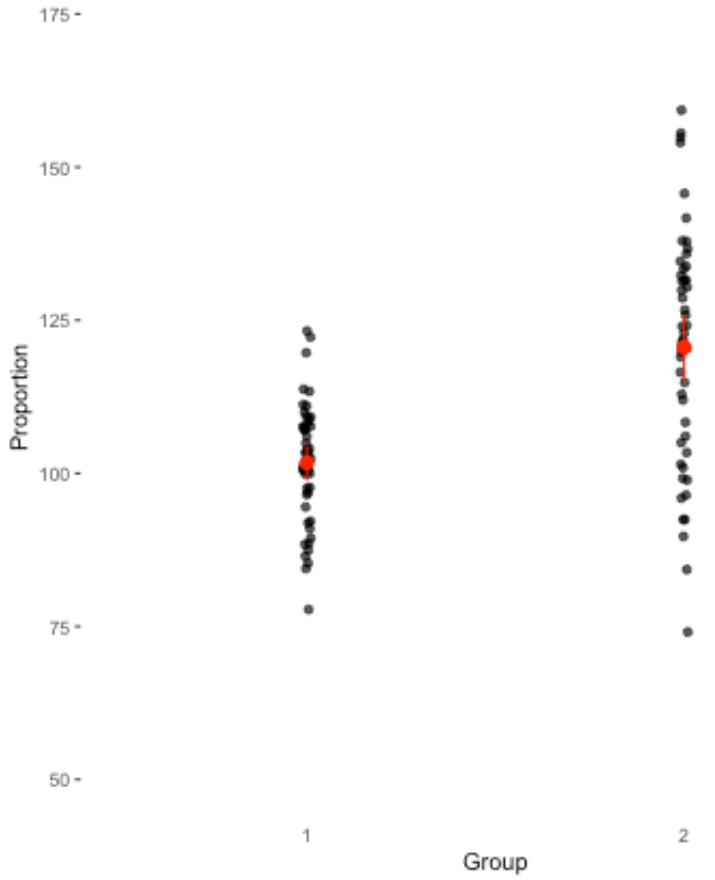
We need some data

```
group_1 <- rnorm(50, 100, 10)
group_2 <- rnorm(50, 120, 20)
groups <- rep(c(1,2), each = 50)
data_groups <- tibble(IQ = c(group_1, group_2),
group = groups)
```

Compare two groups

```
ggplot(data_groups,
       aes(x = factor(groups),
           y = IQ)) +
  geom_jitter(width = .01, alpha = .7) +
  stat_summary(fun.data = "mean_cl_boot",
               size = .5, colour = "red")+
  ylim(c(50,175)) +
  theme_bw() +
  labs(y = "Proportion", x = "Group") +
  theme(panel.grid = element_blank(),
        panel.border = element_blank(),
        axis.ticks.x = element_blank())
```

Compare two groups



What did that code do?

Bi-variate correlations

We need some data

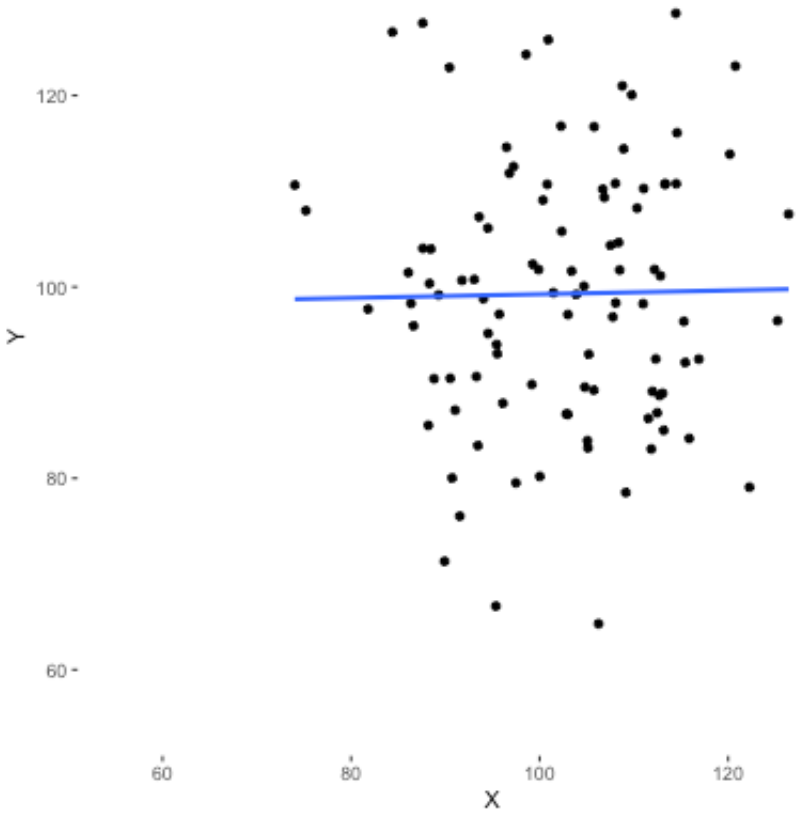
```
x <- rnorm(100, 100, 10)
y <- rnorm(100, 100,14)

corr_data <- tibble(x = x, y = y)
```

Bi-variate correlation

```
ggplot(corr_data, aes(x = x, y = y)) +
  geom_point()+
  geom_smooth(method = "lm",
              se = FALSE)+
  ylim(c(55,135))+
  xlim(c(55,135))+
  theme_bw() +
  labs(y = "Y", x = "X") +
  theme(panel.grid = element_blank(),
        panel.border = element_blank())
```

Bi-variate correlation



What did that code do?

ANOVA (like) designs

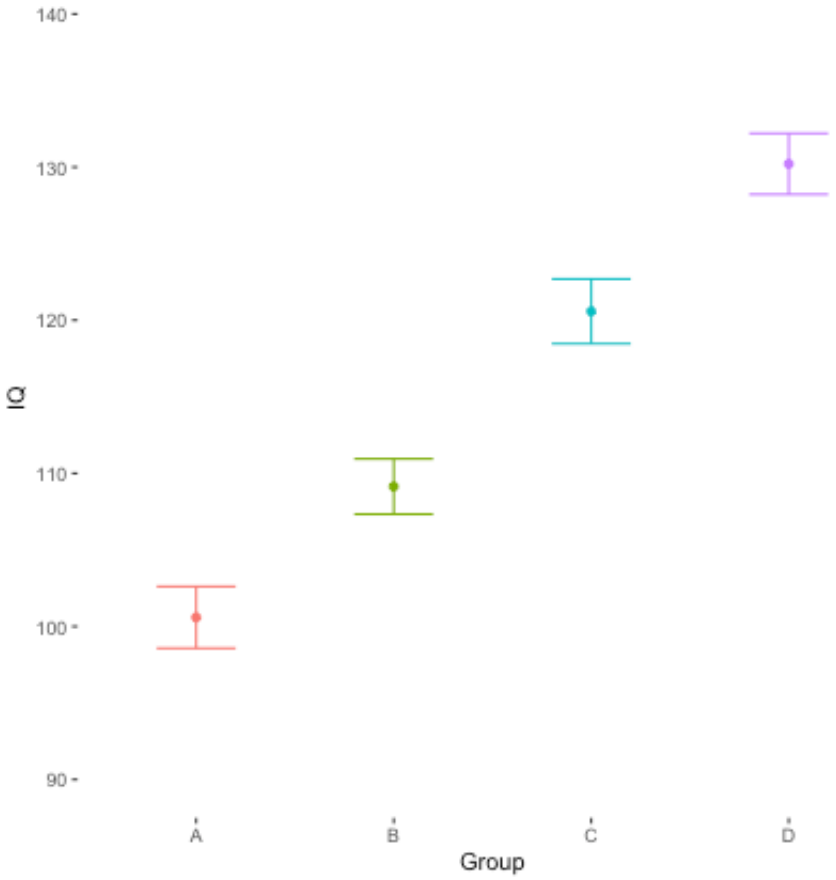
We need some data

```
a_data <- rnorm(100, 100, 10)
b_data <- rnorm(100, 110, 10)
c_data <- rnorm(100, 120, 10)
d_data <- rnorm(100, 130, 10)
aov_data <- tibble(dat =
                     c(a_data, b_data,
                       c_data, d_data),
                   group = factor(rep(c("A", "B",
                                        "C", "D") ,
                                      each = 100)))
```

One factor ANOVA

```
aov_data %>%
  group_by(group) %>%
  summarise(mean = mean(dat, na.rm = TRUE),
            se = sd(dat)/sqrt(sum(!is.na(dat)))) %>%
  ggplot(aes(x = group, y = mean,
             color = group)) +
  geom_point() +
  geom_errorbar(aes(ymin = mean - 1.96*se,
                    ymax = mean + 1.96*se),
                width = .4) +
  ylim(c(90,140)) +
  theme_bw() +
  labs(y = "IQ", x = "Group") +
  theme(panel.grid = element_blank(),
        panel.border = element_blank()) +
  guides(color = "none")
```

One factor ANOVA



What did that code do?

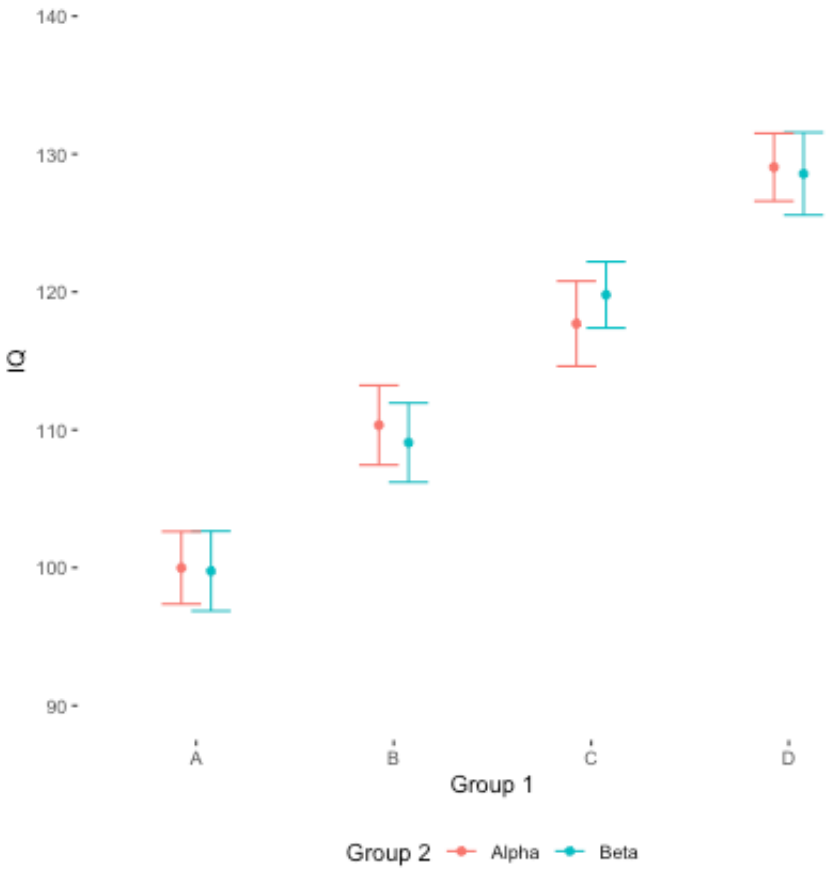
We need some more data

```
a_data <- rnorm(100, 100, 10)
b_data <- rnorm(100, 110, 10)
c_data <- rnorm(100, 120, 10)
d_data <- rnorm(100, 130, 10)
aov_data <- tibble(dat = c(a_data, b_data,
                           c_data, d_data),
                   group_1 = factor(rep(c("A", "B",
                                          "C", "D"),
                                        each = 100)),
                   group_2 = factor(rep(rep(c("Alpha", "Beta"),
                                            each = 50),4)))
```

Two factor ANOVA (like) design

```
aov_data %>%
  group_by(group_1, group_2) %>%
  summarise(mean = mean(dat,
                        na.rm = TRUE),
            se = sd(dat)/sqrt(sum(!is.na(dat)))) %>%
  ggplot(aes(x = group_1, y = mean,
             color = group_2)) +
  geom_point(position = position_dodge(width = .3)) +
  geom_errorbar(aes(ymin = mean - 1.96*se,
                    ymax = mean + 1.96*se), width = .4,
                position = position_dodge(width = .3)) +
  ylim(c(90,140))+
  theme_bw() +
  labs(y = "IQ", x = "Group 1") +
  theme(panel.grid = element_blank(),
        panel.border = element_blank(),
        legend.position = "bottom") +
  scale_color_discrete(name = "Group 2")
```

Two factor ANOVA (like) design



What did that code do?

Graph for multiple regression

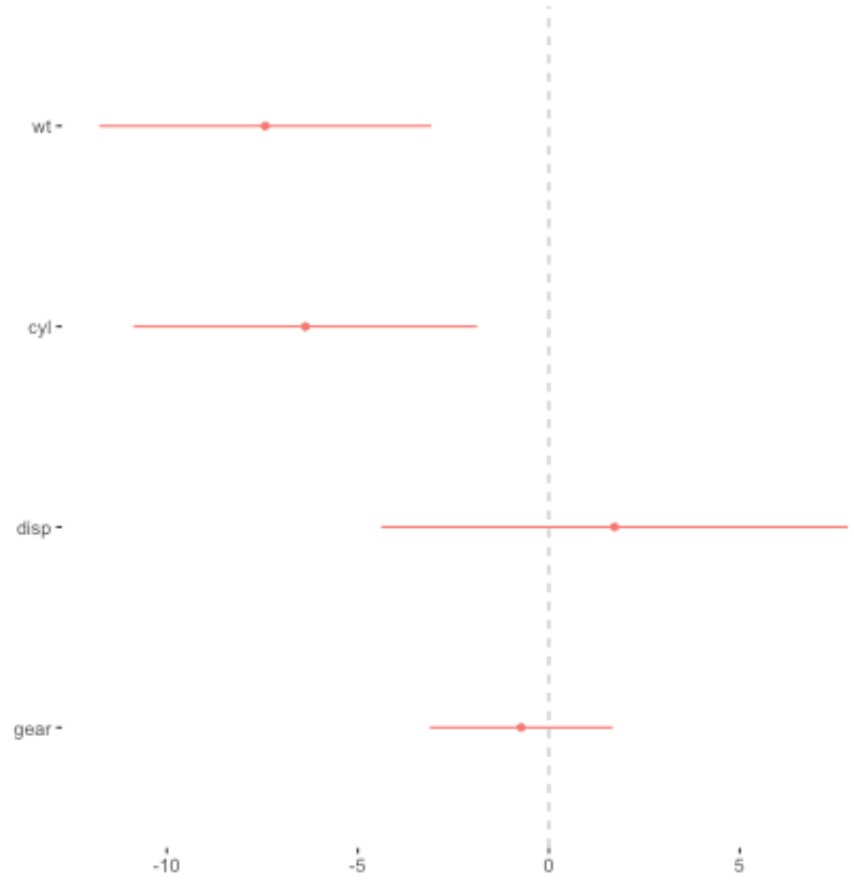
The code

```
library(dotwhisker)

m1 <- lm(mpg ~ wt + cyl + disp + gear, data = mtcars)

dwplot(m1) +
  theme_bw() +
  theme(panel.grid = element_blank(),
        panel.border = element_blank(),
        legend.position = "none") +
  geom_vline(xintercept = 0, color = "grey", linetype = "dashed")
```

The graph



Examination

- The examination for this week is optional.

1. Choose two graphs from published papers.
2. Recreate the graphs as closely as possible using `ggplot2`. The exercise is most fun if you chose graphs that are not so good.
3. Improve the graphs by removing elements and/or highlighting data to make them more clear.

- If you need to extract data from published graphs you can use the free web tool **WebPlotDigitizer**, which is available at: https://automeris.io/WebPlotDigitizer/

That's all folks!