

Analysis of Correlated Data (CHL5222H) Final Project

Beating the Blues

Meagan Lacroix

```
btb <- read.table(here("data", "btheb.txt"), header = TRUE)
```

```
names(btb)
```

```
[1] "drug"      "length"    "treatment" "bdi.pre"   "bdi.2m"    "bdi.4m"
[7] "bdi.6m"    "bdi.8m"
```

```
head(btb)
```

	drug	length	treatment	bdi.pre	bdi.2m	bdi.4m	bdi.6m	bdi.8m
1	No	>6m	TAU	29	2	2	NA	NA
2	Yes	>6m	BtheB	32	16	24	17	20
3	Yes	<6m	TAU	25	20	NA	NA	NA
4	No	>6m	BtheB	21	17	16	10	9
5	Yes	>6m	BtheB	26	23	NA	NA	NA
6	Yes	<6m	BtheB	7	0	0	0	0

```
tail(btb)
```

	drug	length	treatment	bdi.pre	bdi.2m	bdi.4m	bdi.6m	bdi.8m
95	No	>6m	BtheB	16	11	4	2	3
96	Yes	>6m	BtheB	16	16	10	10	8
97	Yes	<6m	TAU	28	NA	NA	NA	NA
98	No	>6m	BtheB	11	22	9	11	11
99	No	<6m	TAU	13	5	5	0	6
100	Yes	<6m	TAU	43	NA	NA	NA	NA

```
#add participant ID
btb <- btb %>%
  mutate(id = row_number()) %>%
  select(id, everything())
```

```
btb <- btb %>%
  mutate(
    drug = factor(drug, levels = c("No", "Yes")), # Set reference level
    length = factor(length, levels = c("<6m", ">6m")), # Set reference level
    treatment = factor(treatment, levels = c("TAU", "BtheB")) # Set TAU as reference
  )
```

```
str(btb) # Check structure
```

```
'data.frame': 100 obs. of 9 variables:
 $ id      : int  1 2 3 4 5 6 7 8 9 10 ...
 $ drug    : Factor w/ 2 levels "No","Yes": 1 2 2 1 2 2 2 1 2 2 ...
 $ length  : Factor w/ 2 levels "<6m",">6m": 2 2 1 2 2 1 1 2 1 2 ...
 $ treatment: Factor w/ 2 levels "TAU","BtheB": 1 2 1 2 2 2 2 1 1 2 2 ...
 $ bdi.pre : int  29 32 25 21 26 7 17 20 18 20 ...
 $ bdi.2m  : int  2 16 20 17 23 0 7 20 13 5 ...
 $ bdi.4m  : int  2 24 NA 16 NA 0 7 21 14 5 ...
 $ bdi.6m  : int  NA 17 NA 10 NA 0 3 19 20 8 ...
 $ bdi.8m  : int  NA 20 NA 9 NA 0 7 13 11 12 ...
```

```
table(btb$drug) # See counts for each category
```

```
No Yes
56 44
```

```
table(btb$treatment)
```

```
TAU BtheB
48 52
```

```
#make data set long
btb.long <- btb %>%
pivot_longer(cols = c("bdi.pre", "bdi.2m", "bdi.4m", "bdi.6m", "bdi.8m"),
names_to = "visit",
values_to = "score") %>%
  mutate(
    occasion = case_when(
      visit == "bdi.pre" ~ 0,
      visit == "bdi.2m" ~ 1,
      visit == "bdi.4m" ~ 2,
      visit == "bdi.6m" ~ 3,
      visit == "bdi.8m" ~ 4))
```

```
btb_summary <- btb.long %>%
  group_by(occasion, treatment, drug, length) %>%
  summarise(
    mean_bdi = mean(score, na.rm = TRUE),
    sd_bdi = sd(score, na.rm = TRUE),
    lower_ci = mean_bdi - 1.96 * sd_bdi,
    upper_ci = mean_bdi + 1.96 * sd_bdi
  ) %>%
  ungroup()
```

```
ggplot(btb_summary, aes(x = occasion, y = mean_bdi, color = treatment, group = treatment)) +
  geom_line(size = 1) + # Line plot for mean BDI over time
  geom_point(size = 2) + # Add points for means
  geom_errorbar(aes(ymin = lower_ci, ymax = upper_ci), width = 0.2) + # Add 95% CI
  facet_grid(drug ~ length) + # Panel by drug and length groups
  labs(title = "Mean BDI Score Over Time by Treatment",
       x = "Time (Occasion)", y = "Mean BDI Score",
       color = "Treatment Group") +
  theme_minimal() +
  theme(strip.text = element_text(size = 12, face = "bold"))
```

Mean BDI Score Over Time by Treatment

