Backround/Context

The data we are using is from a clinical trial called "Beat the Blues". The Beat the Blues computer program was designed to deliver cognative behavioural therapy to depressed patients. This dataset is from the R package "HSAUR2". The data contains 200 observations each relating to a study subject. The study measured 8 variables: drug, length, treatment, bdi.pre, bdi.2m, bdi.3m, bdi.5m and bdi.8m. The variable "drug" refers to if the subject took anti-depressants, "length" is the length of the subjects current episode of depression (either less than 6 months or more than six months), "treatment" is the treatment the subject was placed on (either TAU, treatment as usual, or BtheB, Beat the Blues), the bdi variables refer to the Beck Depression Inventory at certain points during the study.

```
library("HSAUR2")

## Loading required package: tools
library("ggplot2")
attach(BtheB)
```

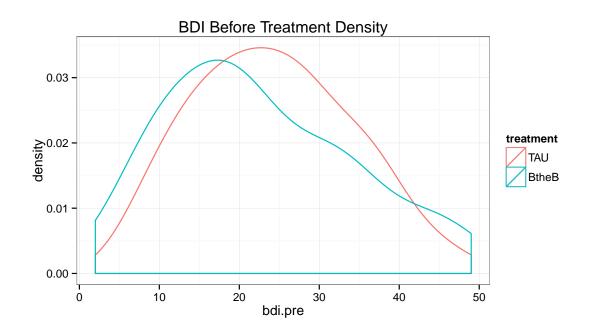
```
summary(BtheB[, c(1:4, 6, 8)])
##
     drug
             length
                       treatment
                                     bdi.pre
                                                      bdi.3m
                                                                      bdi.8m
##
    No :56
             <6m:49
                       TAU :48
                                  Min.
                                          : 2.0
                                                  Min.
                                                          : 0.0
                                                                          : 0.0
                                                                  Min.
    Yes:44
             >6m:51
                       BtheB:52
                                  1st Qu.:15.0
##
                                                  1st Qu.: 6.0
                                                                  1st Qu.: 3.0
                                  Median:22.0
##
                                                  Median:13.0
                                                                  Median:10.5
##
                                  Mean
                                          :23.3
                                                  Mean
                                                          :14.8
                                                                  Mean
                                                                          :11.1
##
                                  3rd Qu.:30.2
                                                                  3rd Qu.:15.2
                                                  3rd Qu.:20.0
                                                                  Max.
##
                                          :49.0
                                                  Max.
                                                          :53.0
                                                                          :40.0
                                  Max.
##
                                                  NA's
                                                          :27
                                                                  NA's
                                                                       :48
```

Variable Descriptions

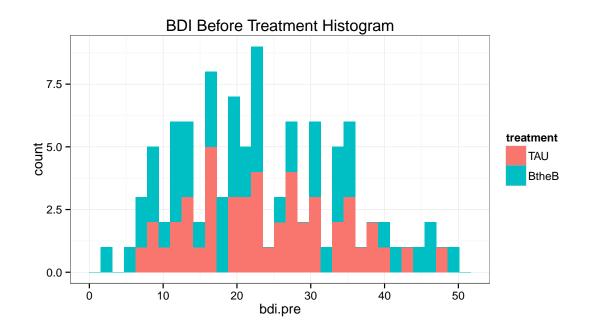
The variables we are interested in are "bdi.pre" and "bdi. 8m", Beck Depression Inventory II before treatment and after six months follow-up.

```
theme_set(theme_bw())
# Subsetting the data
BtheB.tau = BtheB[treatment == "TAU", ]
BtheB.btb = BtheB[treatment == "BtheB", ]
```

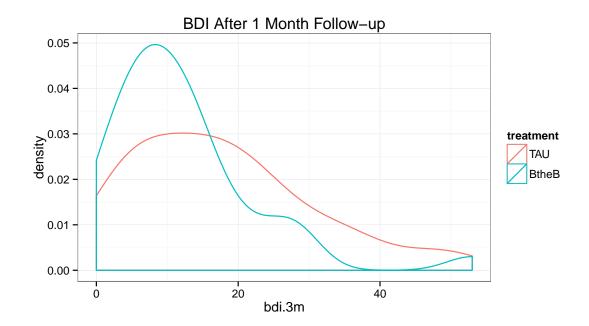
qplot(bdi.pre, geom = "density", color = treatment, main = "BDI Before Treatment Density")



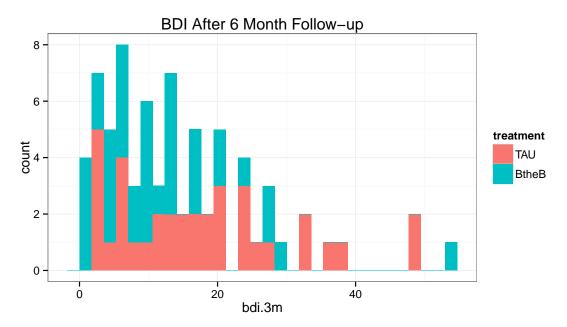
qplot(bdi.pre, fill = treatment, main = "BDI Before Treatment Histogram")
stat_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.



qplot(bdi.3m, geom = "density", color = treatment, main = "BDI After 1 Month Follow-up")



```
qplot(bdi.3m, fill = treatment, main = "BDI After 6 Month Follow-up")
## stat_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.
```



The before treatment groups look similar. However, the after 1 month follow-up treatment groups look severely different. After 1 month follow-up, it seems more BtheB subjects score lower on the BDI scale than the TAU group (which means the subjects are less depressed). This suggests the BtheB treatment works better than treatment as usual.

Missing Data

We chose these two variables to measure the association using a baseline measure and a midpoint measure. Using the 3 month mark (after 1 month follow-up), only 27 subjects having missing data. We originally wished to use the end of the study, however 48 subjects had missing data.

```
sum(is.na(bdi.pre) == TRUE)

## [1] 0

sum(is.na(bdi.3m) == TRUE)

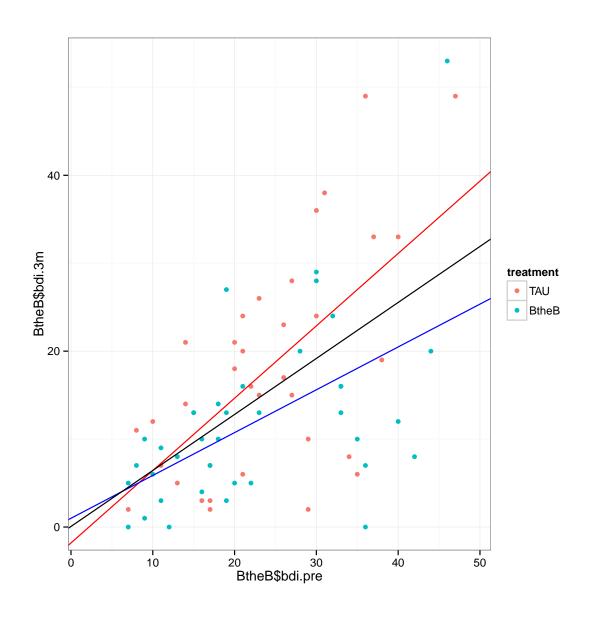
## [1] 27
```

```
sum(is.na(bdi.8m) == TRUE)
## [1] 48
```

Hypothesis

Analysis

```
blm <- lm(bdi.3m ~ bdi.pre)
blm
##
## Call:
## lm(formula = bdi.3m ~ bdi.pre)
## Coefficients:
## (Intercept)
               bdi.pre
##
       0.0645
                   0.6369
blm1 <- lm(bdi.3m ~ bdi.pre, data = BtheB.tau)
blm1
##
## lm(formula = bdi.3m ~ bdi.pre, data = BtheB.tau)
##
## Coefficients:
## (Intercept) bdi.pre
##
       -1.814
                    0.823
blm2 <- lm(bdi.3m ~ bdi.pre, data = BtheB.treat)
## Error: object 'BtheB.treat' not found
b1m2
## Error: object 'blm2' not found
qplot(BtheB$bdi.pre, BtheB$bdi.3m, data = BtheB, color = treatment) + geom_abline(intercept
    slope = 0.8231, colour = "red") + geom_abline(intercept = 0.9949, slope = 0.4871,
    colour = "blue") + geom_abline(intercept = 0.06449, slope = 0.63686, colour = "black")
## Warning: Removed 27 rows containing missing values (geom_point).
```



Results