Regression_Estimate

```r
high_data <-read.csv('../data/highDim_dataset.csv')
low_data <-read.csv('../data/lowDim_dataset.csv')

N_high <- dim(high_data)[1]
N_low <- dim(low_data)[1]

high_data_X <- high_data[,3:dim(high_data)[2]]
low_data_X <- low_data[,3:dim(low_data)[2]]

high_treated <- high_data[high_data$A==1,-2]
high_untreated <- high_data[high_data$A==0,-2]

N_high_treated <- dim(high_treated)[1]
N_high_untreated <- dim(high_untreated)[1]

low_treated <- low_data[low_data$A==1,-2]
low_untreated <- low_data[low_data$A==0,-2]

N_low_treated <- dim(low_treated)[1]
N_low_untreated <- dim(low_untreated)[1]
```

```r
time<- system.time({
  high_treated_lm <- lm(Y~.,data = high_treated);
  high_untreated_lm <- lm(Y~.,data = high_untreated);
  high_treated_predict_all <- predict(high_treated_lm,newdata = high_data_X);
  high_untreated_predict_all <- predict(high_untreated_lm,newdata = high_data_X)})
train_time_high <- time[1]
train_time_high
```

```
## user.self
##     0.164
```

```r
time<- system.time({
  low_treated_lm <- lm(Y~.,data = low_treated);
  low_untreated_lm <- lm(Y~.,data = low_untreated);
  low_treated_predict_all <- predict(low_treated_lm,newdata = low_data_X);
  low_untreated_predict_all <- predict(low_untreated_lm,newdata = low_data_X)})
train_time_low <- time[1]
train_time_low
```

```
## user.self
##      0.01
```

```r
reg_est_ATE_high<-sum(high_treated_predict_all - high_untreated_predict_all)/N_high
reg_est_ATE_low<-sum(low_treated_predict_all - low_untreated_predict_all)/N_low
```

```r
reg_est_ATE_high
```

```
## [1] -2.95978
```

```r
reg_est_ATE_low
```

```
## [1] 2.526944
```

```
# True ATE:
true_ATE_high <- -3
true_ATE_low <- 2.5

# Comparison:

true_ATE_high - reg_est_ATE_high
```

```
## [1] -0.04022036
```

```
true_ATE_low - reg_est_ATE_low
```

```
## [1] -0.02694398
```