Task4

zhexuan

2025-10-06

Task 4: The Follow-up Study of DiGeMon-123 in Rats

In this task, we model the data to investigate if the DiGeMon-123 has any effect on Lee index, and present and discuss our findings.

Load Data

We first present the data4. It includes 73 observations, each representing one rat, with information on treatment group (TRT), treatment duration (Time, in weeks), and Lee Index (LeeIdx). The scatter plot and box plot of the raw data are plotted below.

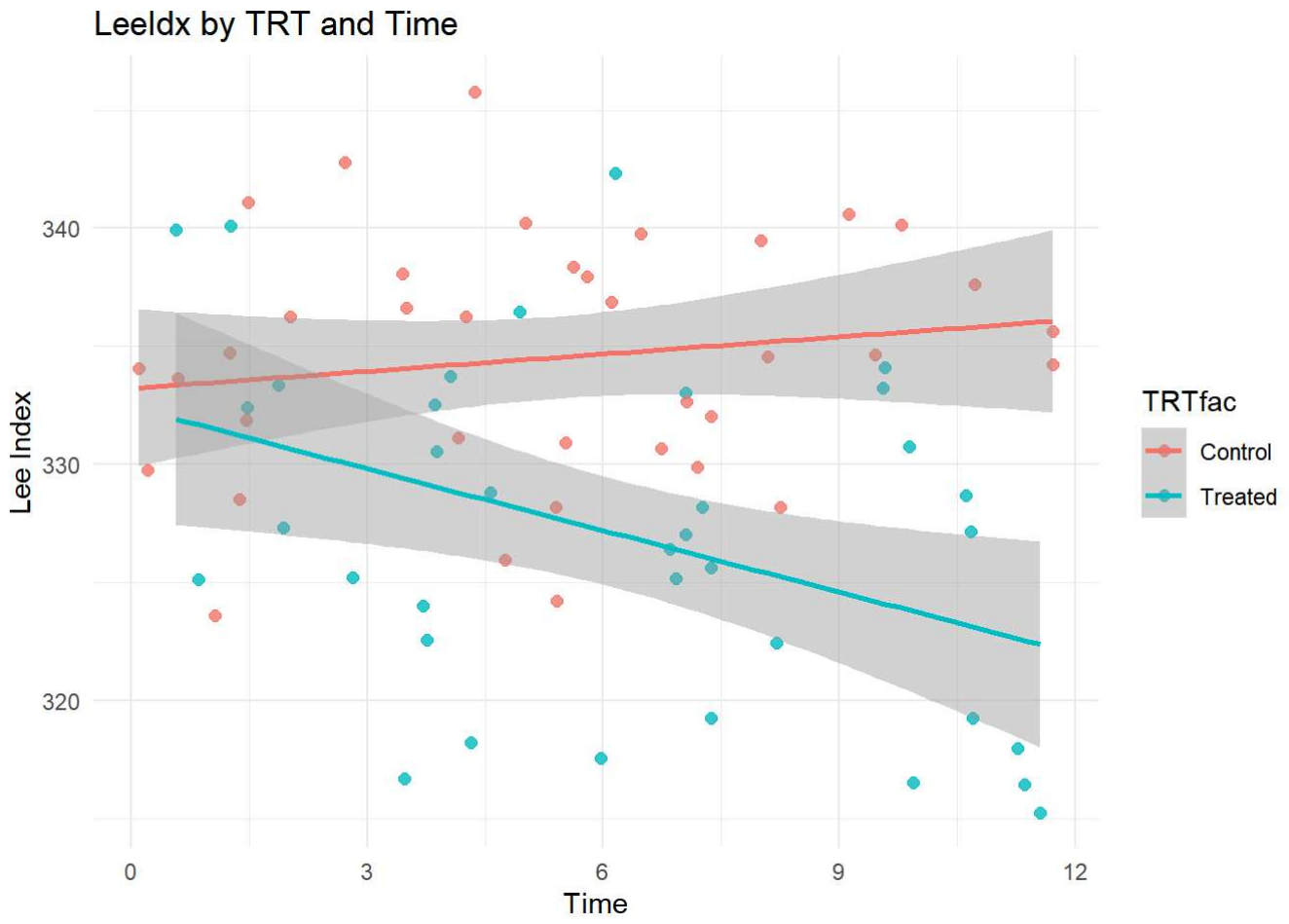
```
df <- read.csv("Data_T4.csv", stringsAsFactors = FALSE)
names(df) <- make.names(names(df))
df <- df %>% rename(ID = X, LeeIdx = LeeIdx, TRT = TRT, Time = Time)
df$ID <- as.factor(df$ID)
df$TRT <- as.numeric(df$TRT)   # 0/1 numeric
df$TRTfac <- factor(df$TRT, levels = c(0,1), labels = c("Control","Treated"))


print(df %>% group_by(TRTfac) %>% summarise(n = n(), mean_Lee = mean(LeeIdx, na.rm=TRUE), sd_Le
e = sd(LeeIdx, na.rm=TRUE)))
```

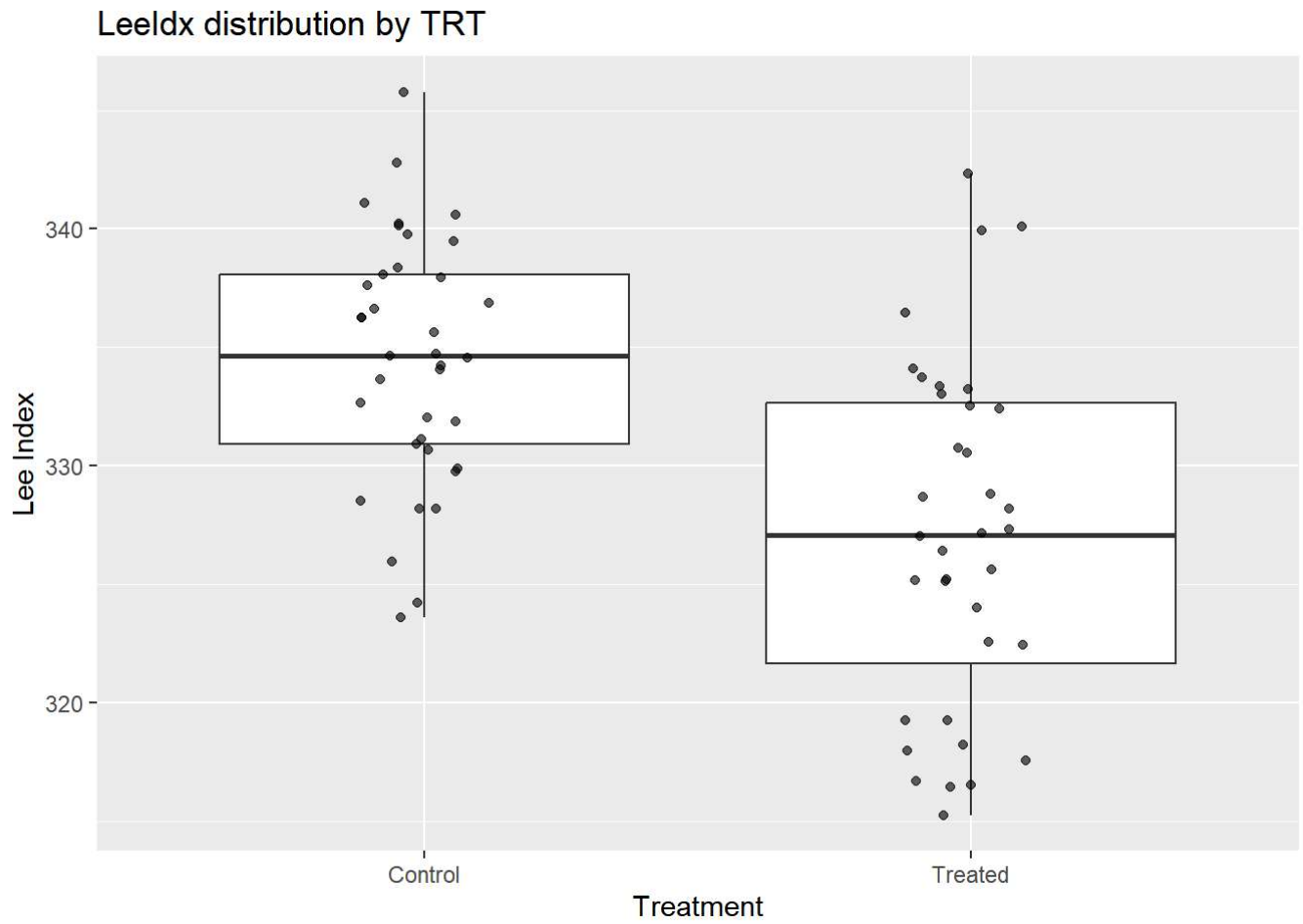
```
## # A tibble: 2 × 4
##   TRTfac      n mean_Lee sd_Lee
##   <fct>   <int>    <dbl>  <dbl>
## 1 Control    37     335.   5.20
## 2 Treated    36     327.   7.30
```

```
p1 <- ggplot(df, aes(x = Time, y = LeeIdx, color = TRTfac)) +
  geom_point(size = 2, alpha = 0.8) +
  geom_smooth(method = "lm", se = TRUE) +
  labs(title = "LeeIdx by TRT and Time", x = "Time", y = "Lee Index") +
  theme_minimal()
print(p1)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
# scatter plot & box plot
p2 <- ggplot(df, aes(x = TRTfac, y = LeeIdx)) +
  geom_boxplot() + geom_jitter(width = 0.12, alpha = 0.6) +
  labs(title = "LeeIdx distribution by TRT", x = "Treatment", y = "Lee Index")
print(p2)
```



Model selection & Comparison

Then we need to choose appropriate models to fit the data. Since each individual was only measured once, LeeIdx is a continuous outcome variable, and from the LeeIdx visualization it seems that there is no significant non-linear change over time, a linear regression model is appropriate here. To examine the effects of treatment and time on the Lee Index, two linear models were first fitted. Model 1 included only the main effects of treatment group (TRTfac) and time, while Model 2 included an additional interaction term between treatment and time. The purpose of this comparison was to test whether the rate of change in Lee Index over time differed between the treated and control groups.

Model 1 imposes the assumption that “time has the same effect on both groups”. The treated group had on average a 7.19-unit lower Lee Index than the control group ($p < 0.001$). Time showed a small non-significant negative trend with p-value larger than 0.05, indicating that if groups are not distinguished, the overall change over time is not significant.

From the results of model 2, neither the main effect of treatment nor that of time was statistically significant on their own (both $p > 0.05$), suggesting that at the baseline level, neither factor alone could explain much variation in the Lee Index. However, the interaction term (TRTfac:Time) was significant ($p < 0.05$). This indicates that the effect of time on the Lee Index depends on the treatment group — in other words, the slopes of the two lines differ. Specifically, the treated group shows a steeper negative slope, implying that their Lee Index decreases faster over time compared to the control group. It decreased by approximately 1.11 per unit time.

To formally compare the two models, an ANOVA was conducted between Model 1 and Model 2. The result ($F = 6.65$, $p = 0.012$) shows that including the interaction term significantly improves model fit, confirming that the interaction effect is meaningful and should be retained. Thus, Model 2 was used for all subsequent analyses.

```
mod1 <- lm(LeeIdx ~ TRTfac + Time, data = df)
summary(mod1)
```

```
##
## Call:
## lm(formula = LeeIdx ~ TRTfac + Time, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.3086  -3.9353   0.6504   4.8526  15.2958
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   336.2384     1.5792 212.919  < 2e-16 ***
## TRTfacTreated  -7.1945     1.4812  -4.857 7.01e-06 ***
## Time           -0.3238     0.2241  -1.445    0.153
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.274 on 70 degrees of freedom
## Multiple R-squared:  0.2854, Adjusted R-squared:  0.265
## F-statistic: 13.98 on 2 and 70 DF,  p-value: 7.803e-06
```

```
# with interaction (whether the therapeutic effect changes over time)
mod2 <- lm(LeeIdx ~ TRTfac * Time, data = df)
summary(mod2)
```

```
##
## Call:
## lm(formula = LeeIdx ~ TRTfac * Time, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.6969  -4.6524   0.3027   4.0205  15.2771
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        333.2155     1.9186 173.676   <2e-16 ***
## TRTfacTreated       -0.7944     2.8613  -0.278    0.782
## Time                 0.2428     0.3078   0.789    0.433
## TRTfacTreated:Time  -1.1125     0.4313  -2.579    0.012 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.035 on 69 degrees of freedom
## Multiple R-squared:  0.3482, Adjusted R-squared:  0.3199
## F-statistic: 12.29 on 3 and 69 DF,  p-value: 1.563e-06
```

```
# Compare the two models (whether interaction is needed)
anova_mods <- anova(mod1, mod2)
print(anova_mods)
```

```
## Analysis of Variance Table
##
## Model 1: LeeIdx ~ TRTfac + Time
## Model 2: LeeIdx ~ TRTfac * Time
##   Res.Df    RSS Df Sum of Sq      F  Pr(>F)
## 1     70 2755.5
## 2     69 2513.2  1    242.31 6.6528 0.01203 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Heteroscedasticity SE

Because the presence of heteroscedasticity (unequal variance of residuals) could affect the reliability of standard errors, heteroskedasticity-robust standard errors were computed for Model 2. Using robust SE helps ensure that statistical inference remains valid even if the residual variance is not constant across observations. The significance pattern remained consistent with the original model: the interaction term remained significant, while the main effects did not. This consistency strengthens confidence in the robustness of the interaction effect.

```
#  (heteroskedasticity-robust SE)
cov_hc <- vcovHC(mod2, type = "HC3")
robust_t <- coeftest(mod2, vcov. = cov_hc)
print(robust_t)
```

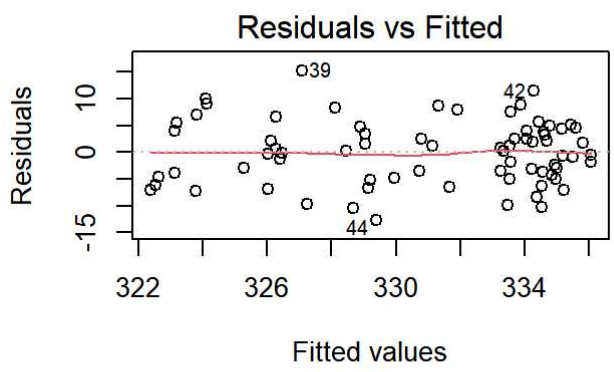
```
##
## t test of coefficients:
##
##                     Estimate Std. Error  t value  Pr(>|t|)
## (Intercept)        333.21546    1.64514 202.5460 < 2.2e-16 ***
## TRTfacTreated       -0.79442    2.95471  -0.2689  0.788835
## Time                 0.24281    0.21769   1.1154  0.268557
## TRTfacTreated:Time  -1.11246    0.40621  -2.7387  0.007843 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

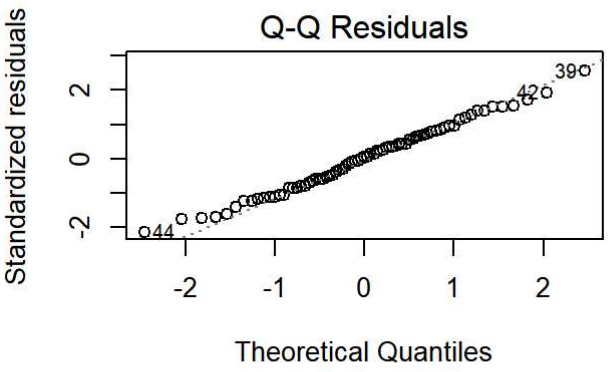
Analysis of Model 2

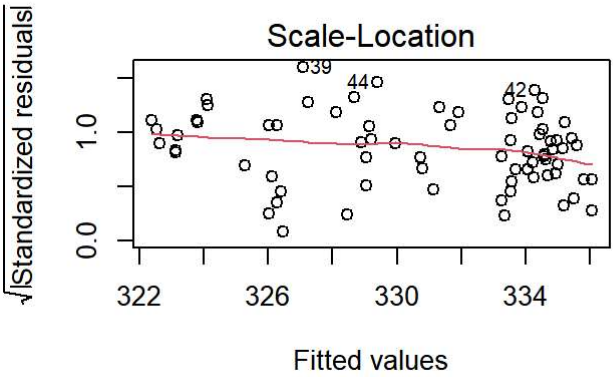
Model diagnostics were then performed to examine assumptions of linear regression. Four plots were generated: residuals vs fitted values, Q-Q plot, scale-location plot, and residuals vs leverage. The residual-fitted plot showed no clear pattern, suggesting approximate linearity and homoscedasticity. The Q-Q plot indicated that most residuals followed a normal distribution, with only minor deviations at the tails. The scale-location plot confirmed relatively constant variance, and the leverage plot revealed no extreme influential points. Overall, these diagnostics supported the adequacy of Model 2.

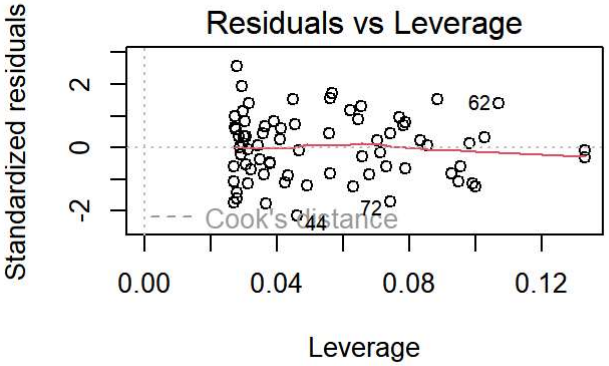
To further confirm residual normality, the Shapiro–Wilk test was conducted. The result ($p > 0.05$) indicated no significant departure from normality. The Breusch–Pagan test (BP = 3.97, df = 3, $p = 0.26$) was then used to assess heteroscedasticity. Since the p-value was greater than 0.05, the null hypothesis of homoscedasticity could not be rejected, suggesting that the variance of residuals was relatively constant across fitted values.

```
# includes Residuals vs Fitted, QQ, Scale-Location, Cook's
par(mfrow=c(2,2))
plot(mod2)
```










```
par(mfrow=c(1,1))

# normality test (Shapiro-Wilk on residuals)
sh_res <- shapiro.test(residuals(mod2))
print(sh_res)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(mod2)
## W = 0.99151, p-value = 0.9098
```

```
# test for heteroscedasticity (Breusch-Pagan)
bp <- bptest(mod2)
print(bp)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  mod2
## BP = 3.9726, df = 3, p-value = 0.2644
```

To interpret the treatment effects more clearly, predicted means of the Lee Index were calculated from Model 2 for each combination of treatment and time, followed by pairwise comparisons between groups. The results revealed that, although baseline differences between groups were small, the treated group demonstrated a

larger decline in the predicted Lee Index over time compared to the control group. This finding quantitatively supports the earlier observation from the significant interaction term.

```
# choose times of practical interest:
times_to_eval <- c(min(df$Time, na.rm=TRUE),
                   median(df$Time, na.rm=TRUE),
                   max(df$Time, na.rm=TRUE))
times_to_eval
```

```
## [1]  0.1063677  5.6242958 11.7129661
```

```
# get emmeans at those times and pairwise contrasts
library(emmeans)
emm_at_times <- emmeans(mod2, ~ TRTfac | Time, at = list(Time = times_to_eval))
emm_at_times
```

```
## Time =  0.106:
##  TRTfac  emmean    SE df lower.CL upper.CL
##  Control    333 1.890 69      329      337
##  Treated    332 2.090 69      328      337
##
## Time =  5.624:
##  TRTfac  emmean    SE df lower.CL upper.CL
##  Control    335 0.996 69      333      337
##  Treated    328 1.020 69      325      330
##
## Time = 11.713:
##  TRTfac  emmean    SE df lower.CL upper.CL
##  Control    336 2.200 69      332      340
##  Treated    322 1.950 69      318      326
##
## Confidence level used: 0.95
```

```
contrast_at_times <- contrast(emm_at_times, method = "pairwise")   # default: treated - control
summary(contrast_at_times, infer = c(TRUE, TRUE))
```

```
## Time =  0.106:
##  contrast          estimate   SE df lower.CL upper.CL t.ratio p.value
##  Control - Treated    0.913 2.82 69    -4.72     6.54   0.323  0.7473
##
## Time =  5.624:
##  contrast          estimate   SE df lower.CL upper.CL t.ratio p.value
##  Control - Treated    7.051 1.43 69     4.21     9.90   4.945  <.0001
##
## Time = 11.713:
##  contrast          estimate   SE df lower.CL upper.CL t.ratio p.value
##  Control - Treated   13.825 2.94 69     7.96    19.69   4.704  <.0001
##
## Confidence level used: 0.95
```

Non-parametric Methods

To validate this conclusion using distribution-free methods, two non-parametric analyses were also performed. Hedges's g was computed to estimate the standardized mean difference between groups. g = 1.169 indicates that there is large treatment effect in reducing Lee Index. In addition, the Wilcoxon rank-sum test was conducted to test whether the median Lee Index differed significantly between treatment groups. And the p-value is very low here, indicating that there was a significant difference in the distribution of LeeIdx between the control group and the treatment group.

Robust Regression Model

Finally, a robust regression model was fitted to further ensure that the results were not driven by outliers or influential data points. Unlike ordinary least squares (OLS), robust regression assigns less weight to extreme values, making parameter estimates more stable. The robust model yielded results similar to Model 2: the interaction term (TRTfac:Time) remained significant, whereas main effects were not. This consistency across models strengthens the reliability of the conclusion that the treatment modifies the rate of change in Lee Index over time.

```
# Cohen's d for TRT groups, unadjusted
cohen_d <- cohen.d(df$LeeIdx ~ df$TRTfac, hedges.correction = TRUE)
print(cohen_d)
```

```
##
## Hedges's g
##
## g estimate: 1.169161 (large)
## 95 percent confidence interval:
##     lower     upper
## 0.6694236 1.6688978
```

```
# Non-parametric comparison(Comparison of the median of the two groups) was used as a robustnes
s test
wilcox_res <- wilcox.test(LeeIdx ~ TRTfac, data = df)
print(wilcox_res)
```

```
##
##  Wilcoxon rank sum exact test
##
## data:  LeeIdx by TRTfac
## W = 1059, p-value = 6.883e-06
## alternative hypothesis: true location shift is not equal to 0
```

```
# robust regression for outliers
rob_mod <- lmrob(LeeIdx ~ TRTfac * Time, data = df)
summary(rob_mod)
```

```
##
## Call:
## lmrob(formula = LeeIdx ~ TRTfac * Time, data = df)
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.6874  -4.2985   0.2877   4.2961  15.4049
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        333.2297     1.5749 211.584  < 2e-16 ***
## TRTfacTreated       -0.6647     3.0411  -0.219  0.82764
## Time                 0.2440     0.2020   1.208  0.23128
## TRTfacTreated:Time  -1.1578     0.4257  -2.720  0.00826 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 6.217
## Multiple R-squared:  0.347,   Adjusted R-squared:  0.3186
## Convergence in 9 IRWLS iterations
##
## Robustness weights:
##  6 weights are ~= 1. The remaining 67 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.5188  0.8871  0.9469  0.9161  0.9844  0.9985
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##         1.548e+00         5.000e-01         4.685e+00         1.000e-07
##           rel.tol         scale.tol         solve.tol          zero.tol
##         1.000e-07         1.000e-10         1.000e-07         1.000e-10
##       eps.outlier             eps.x warn.limit.reject warn.limit.meanrw
##         1.370e-03         2.131e-11         5.000e-01         5.000e-01
##      nResample          max.it       best.r.s       k.fast.s           k.max
##            500              50              2              1            200
##    maxit.scale      trace.lev            mts     compute.rd fast.s.large.n
##            200              0           1000              0           2000
##                   psi           subsampling                   cov
##             "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
```

Conclusion

In summary, the linear model analysis, diagnostic checks, non-parametric tests, and robust regression all lead to a consistent conclusion. The treatment group exhibited a faster decline in Lee Index over time compared to the control group. And the robustness of this finding across multiple analytical approaches provides strong support for the validity of the result.