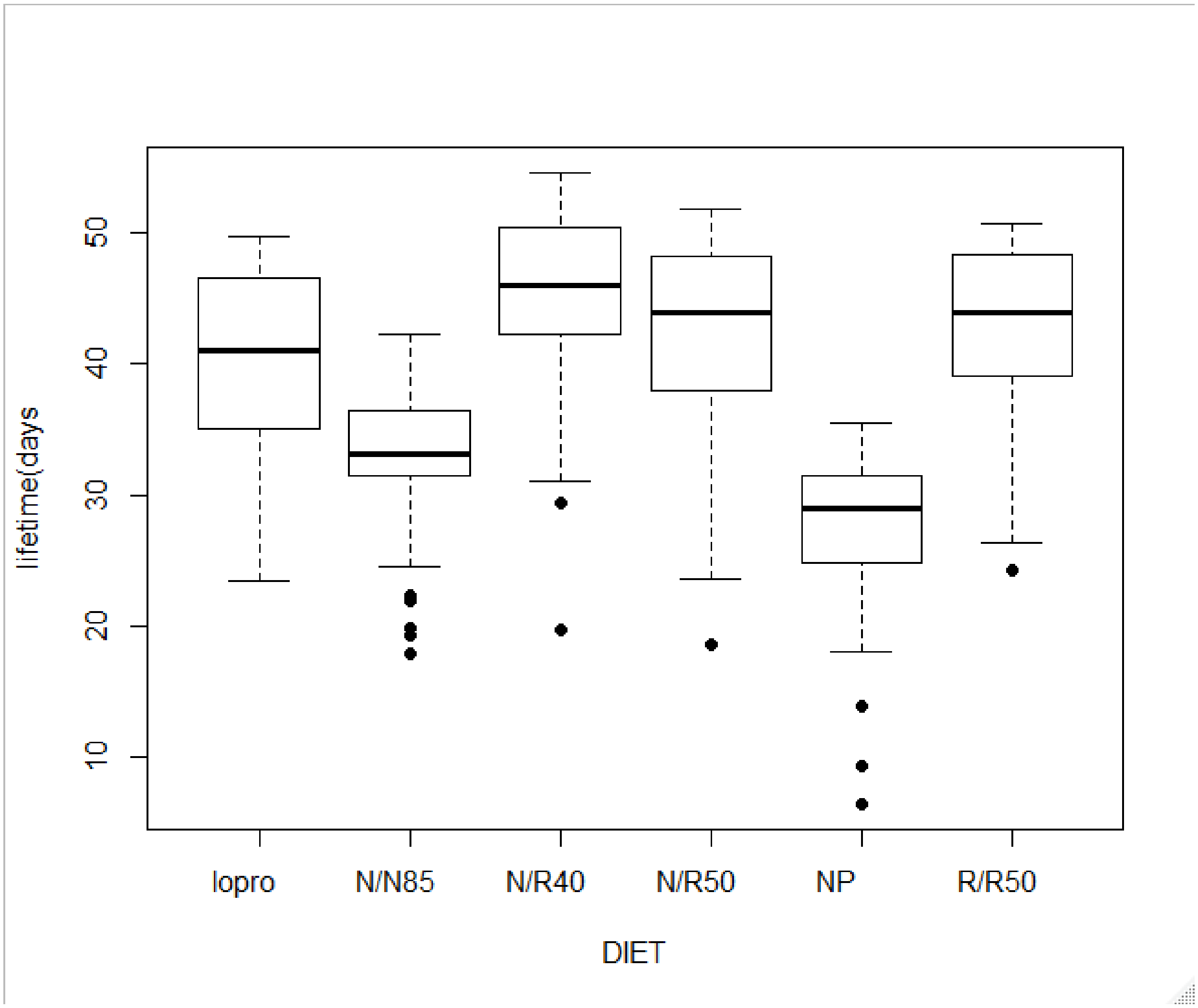
1. Construct a box plot showing lifetime by diet level and discuss the plot. Also calculate the sample means and sample standard deviations for each diet level and discuss your findings.

```
>
> tapply(mice$LIFETIME,mice$DIET,mean)
lopro      N/N85      N/R40      N/R50      NP         R/R50
 39.68571  32.69123  45.11667  42.29718  27.40204  42.88571
> tapply(mice$LIFETIME,mice$DIET,sd)
lopro      N/N85      N/R40      N/R50      NP         R/R50
 6.991694  5.125297  6.703406  7.768195  6.133701  6.683152
>
```

For the sample means, the N/R40 group had the highest mean lifetime, which adds to the hypothesis that a calorie restriction at certain intervals can be beneficial compared to other methods of diets. The NP group had the lowest mean lifetime by quite a large margin. Standard deviations were almost all spread evenly, except for The N/N85 and N/R50 groups with the former having a 5.125 which is lower than the rest and N/R50 having a standard deviation of 7.768 which is higher than the rest of the groups. These differences in standard deviations could be explained by when or if calorie restrictions are implemented for the mice.

The box plot shows that the lopro, N/R40, N/R50, and R/R50 groups had spreads that were approximately similar, although the N/R40 and NP groups had substantial outliers that fell below their groups' respective means. Based on the box plot each group is a little different than the others.



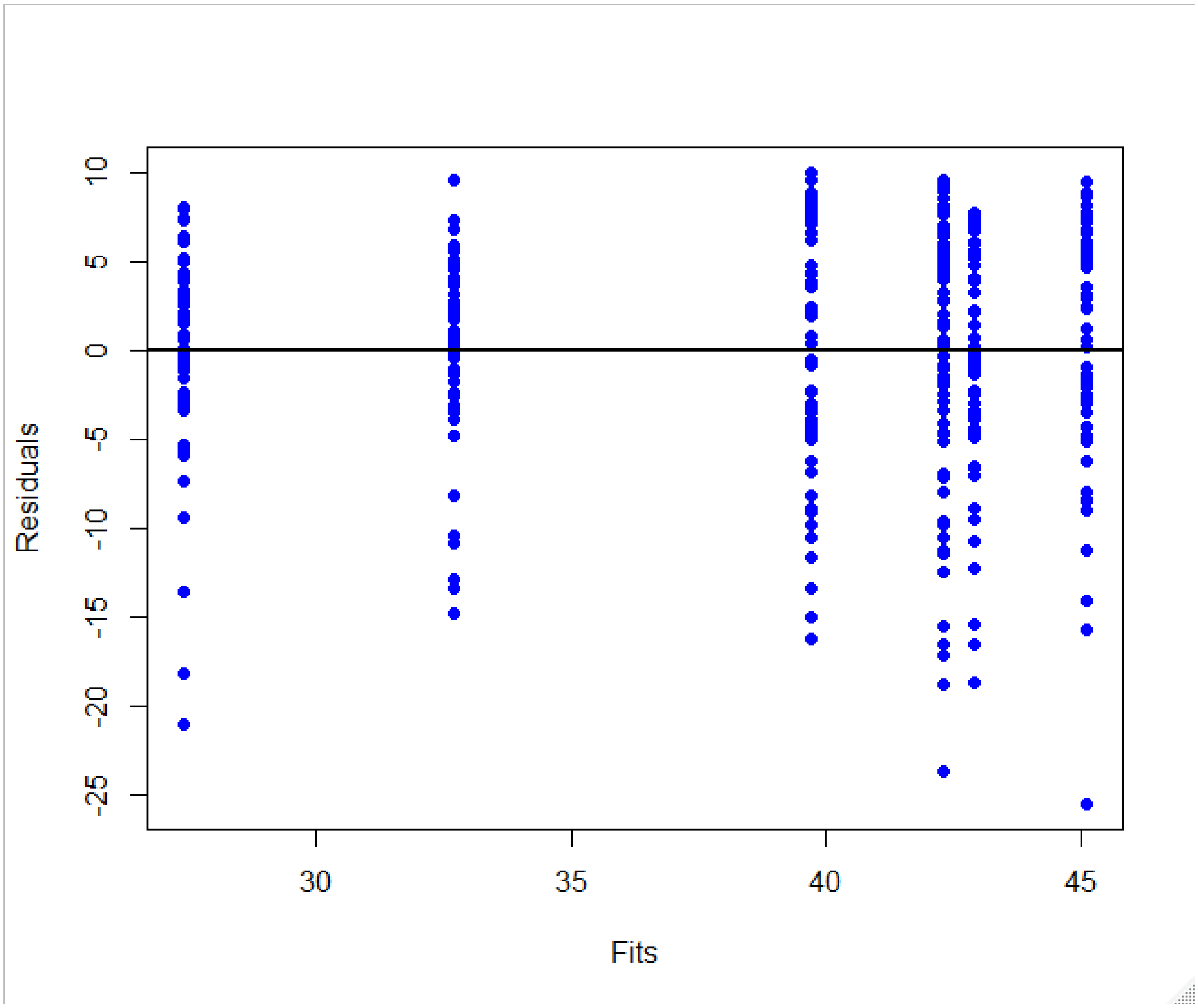
2. Run a One-Way ANOVA and provide the ANOVA table. Discuss your conclusion from the F-test.

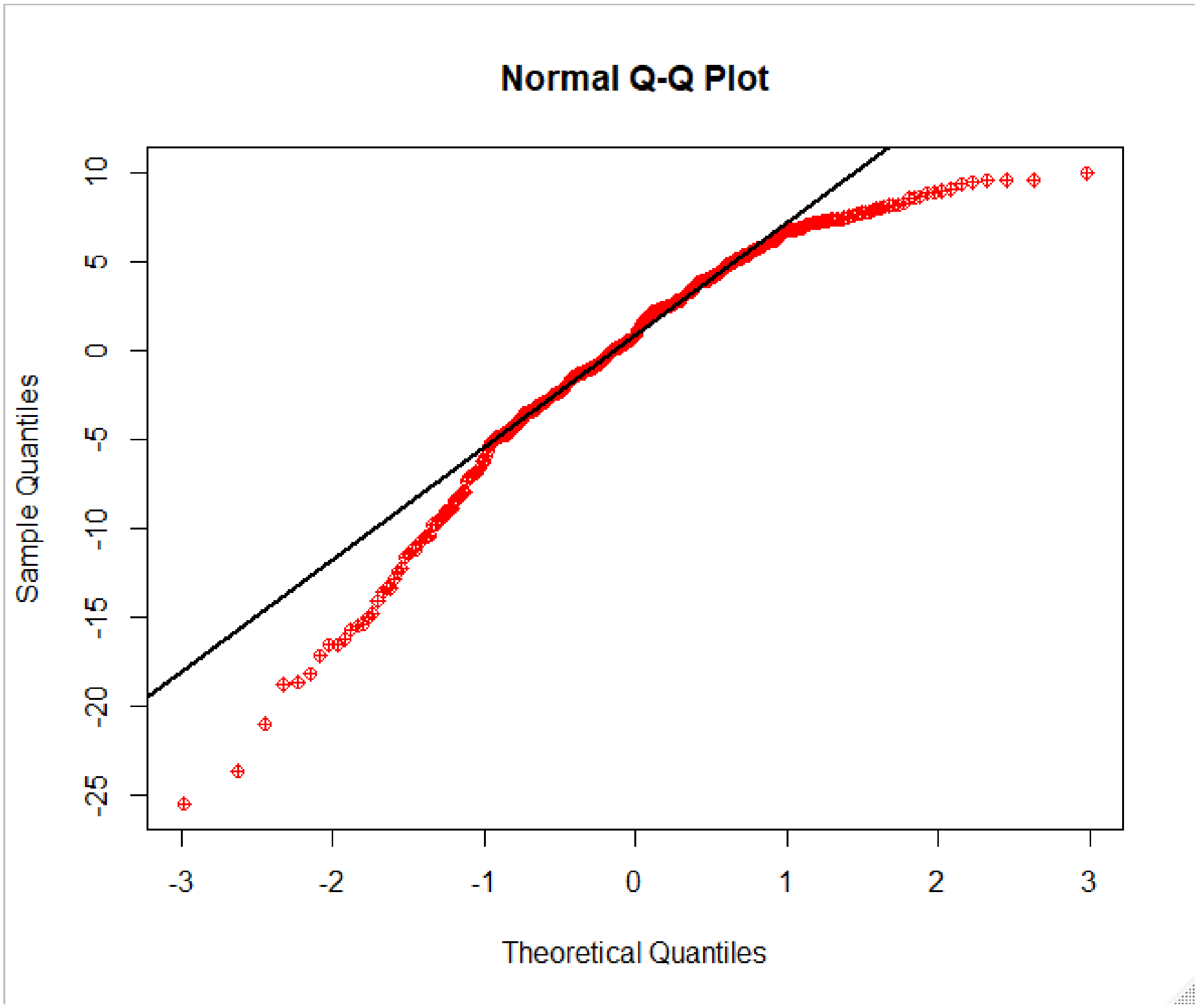
```
Residuals:
     Min       1Q   Median       3Q      Max
-25.5167  -3.3857   0.8143   5.1833  10.0143

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    39.6857     0.8924  44.470  < 2e-16 ***
DIETN/N85      -6.9945     1.2565  -5.567 5.25e-08 ***
DIETN/R40       5.4310     1.2409   4.377 1.60e-05 ***
DIETN/R50       2.6115     1.1936   2.188   0.0293 *
DIETNP        -12.2837     1.3064  -9.403  < 2e-16 ***
DIETR/R50       3.2000     1.2621   2.536   0.0117 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.678 on 343 degrees of freedom
Multiple R-squared:  0.4543,	Adjusted R-squared:  0.4463
F-statistic:  57.1 on 5 and 343 DF,  p-value: < 2.2e-16
```

```
Analysis of Variance Table

Response: LIFETIME
           Df Sum Sq Mean Sq F value    Pr(>F)
DIET        5  12734  2546.8  57.104 < 2.2e-16 ***
Residuals 343  15297    44.6
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
```

Based on the F test/One-way ANOVA table, the diet is very statistically significant, all the diets are marked with significance codes, and the Pr(>F) value is very close to zero denoting its statistical significance.

3. Construct a Residual versus Fits plot (or box plot) as well as a Normal QQ plot.
Describe any weaknesses in the model assumptions.


Residuals
Fits
10
5
0
-5
-10
-15
-20
-25
30
35
40
45



When comparing the residuals versus the fits for equal variance, there seems to be approximately equal variance between them, but when it comes to normality with the QQ plot, there is a departure at the ends of the line. The plot shows that for some of the samples it follows the model assumption of normality but then strays from being considered normal.

4. Construct a pairwise t-test table twice, once with no adjustment for multiple comparisons and once using Bonferroni adjustments. Highlight the significant differences in each table.

```
> pairwise.t.test(mice$LIFETIME,mice$DIET,p.adjust.method = "none")

        Pairwise comparisons using t tests with pooled SD

data:  mice$LIFETIME and mice$DIET

          lopro    N/N85    N/R40    N/R50    NP
N/N85     5.2e-08  -        -        -        -
N/R40     1.6e-05  < 2e-16  -        -        -
N/R50     0.029    1.1e-14  0.017    -        -
NP        < 2e-16  5.9e-05  < 2e-16  < 2e-16  -
R/R50     0.012    8.9e-15  0.073    0.622    < 2e-16

P value adjustment method: none
> pairwise.t.test(mice$LIFETIME,mice$DIET,p.adjust.method = "bonferroni")

        Pairwise comparisons using t tests with pooled SD

data:  mice$LIFETIME and mice$DIET

          lopro    N/N85    N/R40    N/R50    NP
N/N85     7.9e-07  -        -        -        -
N/R40     0.00024  < 2e-16  -        -        -
N/R50     0.44018  1.6e-13  0.24881  -        -
NP        < 2e-16  0.00089  < 2e-16  < 2e-16  -
R/R50     0.17507  1.3e-13  1.00000  1.00000  < 2e-16

P value adjustment method: bonferroni
```

Without adjusting for multiple comparisons, comparing N/R40 and R/R50's t-ratio is closer to zero meaning that there is more similarity between the two before the adjustment because it grows from 0.073 to 1.0. Similarly the comparison between R/R50 and N/R50 grows from 0.622 to 1.0, though for the most part all of the comparison's t-ratios increase after using a Bonferroni adjustment for multiple comparisons.

5. For the comparisons that make sense, discuss your conclusions.

The comparison between R/R50 and N/R50 has a high t-ratio no matter the adjustment method, which means that the groups' differences are less similar than most of the other groups. It seems that the restrictions during or after weaning made a large difference when comparing the groups with the same calorie count.

Comparing NP with lopro, N/R40, and N/R50, it didn't make a difference when adjusting for multiple comparisons. The small t scores signify that these groups are similar with a t-score of 2e-16 in both adjustments.

In conclusion the R/R50, N/R50, and N/R40 (highest) groups had the highest mean lifetimes and the 50 groups were the most similar even though the restrictions came at different points in the lifetime of the mice, which gives suggestive evidence that just reducing calories makes a higher impact on lengthening lifespan