Assignment – use one-way ANOVA to analyze results of a study where 12 participants were randomly assigned to take one of the three types of pain medication (A, B or C) and record times until pain eases off.

Research question – were there any differences in recorded times between the three treatment groups?

First, I prepared the environment for this assignment using the following commands:

```
> rm(list=ls()) #Clear the environment
> setwd("E:/Dropbox/RU DataScience/MSDS660/Week4/Assignment") #Set working directory for the assignme
nt
> getwd() #Check working directory
[1] "E:/Dropbox/RU DataScience/MSDS660/Week4/Assignment"
```

Then, I entered raw data into R using the following commands:

```
> #Input data
>
> #Time to relief (in minutes)
> a <- c(14,24, 12, 25) #Time to relief for A treatment group
> b <- c(20, 14, 17, 18) #time to relief for B treatment group
> c <- c(22, 29, 36, 20) #time to relief for C treatment group
> ttrelief <- c(a, b, c) #combine all time to relief values into one vector
> ttrelief #check the ttrelief content
 [1] 14 24 12 25 20 14 17 18 22 29 36 20
>
> #Treatment levels
>
> k<-3 #number of treatment levels
> n<- 4 #number of observation per treatment level
> #use gl() function to generate factor levels by the pattern of the levels
> treatm <-gl(k,n,12, labels=c("a", "b", "c")) #creating a vector of treatment factors
> treatm #check treatm content
 [1] a a a a b b b b c c c c
Levels: a b c
```

Below is the console window output:

```
E:/Dropbox/RU DataScience/MSDS660/Week4/Assignment/
> getwd() #Check working directory
[1] "E:/Dropbox/RU DataScience/MSDS660/Week4/Assignment"
> #Input data
>
> #Time to relief (in minutes)
> a <- c(14,24, 12, 25) #Time to relief for A treatment group
> b <- c(20, 14, 17, 18) #time to releif for B treatment group
> c <- c(22, 29, 36, 20) #time to releif for C treament group
> ttrelief <- c(a, b, c) #combine all time to releif values into one vector
> ttrelief #check the ttrelief content
 [1] 14 24 12 25 20 14 17 18 22 29 36 20
>
> #Treatment levels
>
> k<-3 #number of treatment levels
> n<- 4 #number of observation per treatment level
> #use gl() function to generate factor levels by the pattern of the levels
> treatm <-gl(k,n,12, labels=c("a", "b", "c")) #creating a vector of treatment factors
> treatm #check treatm content
 [1] a a a a b b b b c c c c
Levels: a b c
```

As a result, all "time to relief" observations are saved as a vector of numerical values ***ttrelief*** and corresponding treatment levels are saved as a vector of factors ***treatm***.

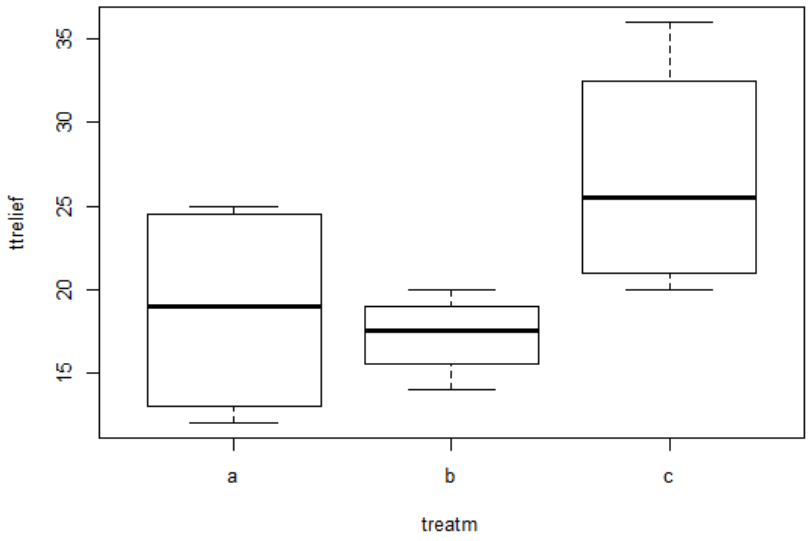
Next step is to explore the data and to check for the ANOVA model assumptions – independence, normality and homogeneity of variance.

According to the assignment, the investigator randomly assigned 12 participants to the treatments, so we can assume that the independence requirement is met for this dataset.

Next, I created side-by-side boxplots for "time to relief" for each treatment group using the following code:

```
#Create side-by-side box plot time to relief for each treatment group
> plot(ttrelief~treatm)
```

The resulting plot is presented below:



I visually inspected the boxplots for outliers, skewness and unequal variance. There are no separated points what would indicated outliers. The boxes for groups A and B are roughly symmetrical, the C box is slightly more extended to show higher values in the third quantile. The A and C boxes are roughly equal in size indicating equal variances, B shows slightly smaller variation however since only four data point were used to construct each box this result can be inconclusive.

Overall, visual analysis in this case did not reveal any obvious problems or counterindications for using the ANOVA model. So, I proceeded with some additional testing.

I used the Shapiro-Wilk test to check for normality of data in each treatment group using the following hypothesis:

Ho: Time to relief for group A is normally distributed

Ha: Time to relief for group A is not normally distributed

This test returned the following results:

```
> shapiro.test(a) #for treatment group A

        Shapiro-Wilk normality test

data:  a
W = 0.83349, p-value = 0.1771
```

The Shapiro-Wilk test returned p-value of 0.1771 for the group A, which is greater than the significance level of 0.05. It means that we can not reject the null hypothesis that the sample came from a normal distribution.

Ho: Time to relief for group B is normally distributed

Ha: Time to relief for group B is not normally distributed

```
> shapiro.test (b) #for treatment group b

        Shapiro-Wilk normality test

data:  b
W = 0.98152, p-value = 0.9109
```

The Shapiro-Wilk test for the group b returned a substantially high p-value of 0.9109, which means that we cannot reject the null hypothesis of normality of the data.

Finally, the hypothesis for the group C are as follow:

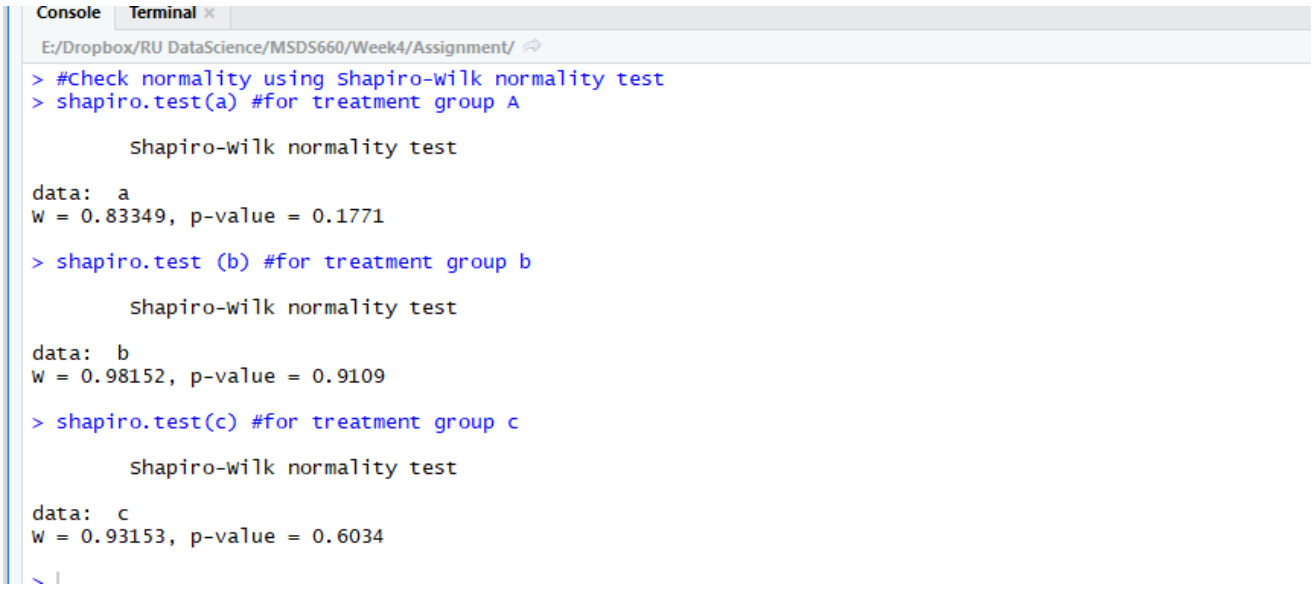
Ho: Time to relief for group C is normally distributed

Ha: Time to relief for group C is not normally distributed

The Shapiro-Wilk test returned the following:

```
> shapiro.test(c) #for treatment group c

        Shapiro-Wilk normality test

data:  c
W = 0.93153, p-value = 0.6034
```

The p-value for the group that was subject to the third treatment (treatment C) is 0.6034 which is considerably higher than the significance level of 0.05. As a result, we can not reject the null hypothesis stating the normality of our data.



To assess homogeneity or approximately equal variability of the data, I used the Levene test (**leveneTest()**) as it is less sensitive to departures from normality. The hypotheses are:

Ho: The variances for all three treatment groups ae equal

Ha: The variances for the three treatment groups are not equal

Below is the console window output.

```
> #use Levene test to check for equality of variance
> library(car)
> leveneTest(ttrelief~treatm) #Levene test with one independent variable
Levene's Test for Homogeneity of Variance (center = median)
      Df F value  Pr(>F)
group  2  4.5444 0.04323 *
       9
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```


```
> #use Levene test to check for equality of variance
> library(car)
> leveneTest(ttrelief~treatm) #Levene test with one independent variable
Levene's Test for Homogeneity of Variance (center = median)
      Df F value  Pr(>F)
group  2  4.5444 0.04323 *
       9
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
```


This test returned a p-value of 0.04 that using the usual α of 0.05 would mean rejecting the null hypothesis. However, considering small sample sizes for each group I adjusted the alpha level to 0.01. With the alpha equal to 0.01 we do not have sufficient proof to reject the null hypothesis of approximate equality of variances.

Conclusion: Overall our analysis of independence, normality and homogeneity of variance did not yield any results showing substantial departure from the ANOVA model assumptions that would prevent us from using it.

So, I proceeded with the analysis. The hypothesis for ANOVA testing are as follows:

Ho: The means for all three treatment groups are equal

Ha: The means for the three treatment groups are not equal

I used ***aov()*** function to fit ANOVA model with the following results:

```
> #fit an analysis of variance model
> avmodel <-  aov(ttrelief~treatm)
> summary(avmodel)
            Df Sum Sq Mean Sq F value Pr(>F)
treatm       2  208.7  104.33   3.007    0.1 .
```

```
Residuals    9  312.2   34.69
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
>
> #fit an analysis of variance model
> avmodel <-  aov(ttrelief~treatm)
> summary(avmodel)
            Df Sum Sq Mean Sq F value Pr(>F)
treatm       2  208.7  104.33   3.007    0.1 .
Residuals    9  312.2   34.69
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
```

According to the above results the F-statistic for the test is 3.007 with the p-value of 0.1 which is higher than the significance level of 0.05. It means that we can not reject the null hypothesis stating that all three means are equal. So, there is no evidence that the treatment had any effect on resulting times to pain relief.

Since aov() function fits an analysis of variance model by calling **lm()** and basically provides a wrapper to lm for fitting linear models, I was interesting in comparing these methods and I decided to check if using lm() directly yields any different results.

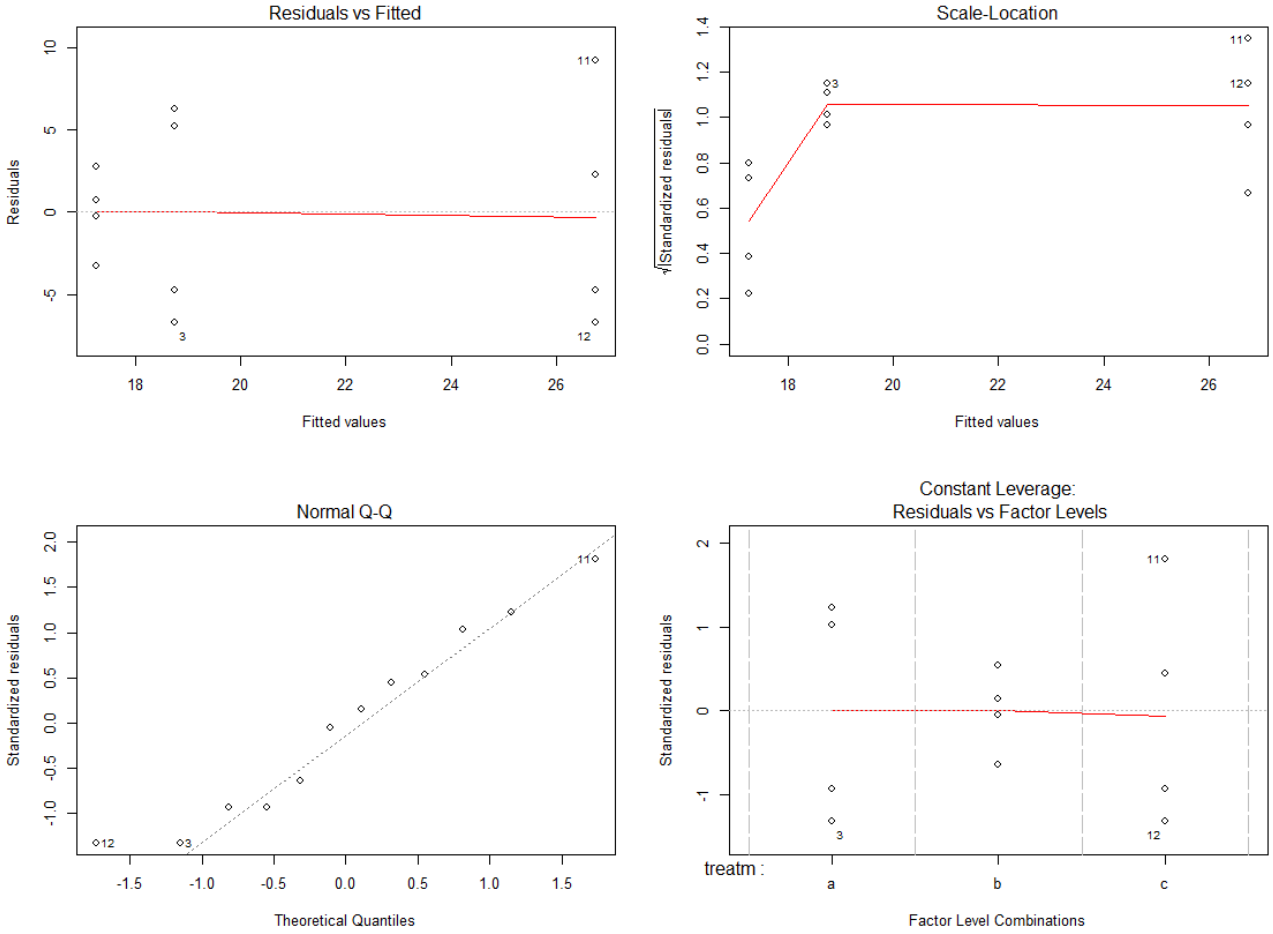
```
> avm <- lm(ttrelief~treatm)
> summary(avm)

Call:
lm(formula = ttrelief ~ treatm)

Residuals:
   Min     1Q Median     3Q    Max
-6.750 -4.750  0.250  3.375  9.250

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   18.750      2.945   6.367  0.00013 ***
treatmb       -1.500      4.165  -0.360  0.72705
treatmc        8.000      4.165   1.921  0.08695 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.89 on 9 degrees of freedom
Multiple R-squared:  0.4006, Adjusted R-squared:  0.2674
F-statistic: 3.007 on 2 and 9 DF,  p-value: 0.09995
```

```
E:/Dropbox/RU DataScience/MSDS660/Week4/Assignment/
> avm <- lm(ttrelief~treatm)
> summary(avm)

Call:
lm(formula = ttrelief ~ treatm)

Residuals:
   Min     1Q Median     3Q    Max
-6.750 -4.750  0.250  3.375  9.250

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   18.750      2.945   6.367  0.00013 ***
treatmb       -1.500      4.165  -0.360  0.72705
treatmc        8.000      4.165   1.921  0.08695 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.89 on 9 degrees of freedom
Multiple R-squared:  0.4006,    Adjusted R-squared:  0.2674
F-statistic: 3.007 on 2 and 9 DF,  p-value: 0.09995

>
```

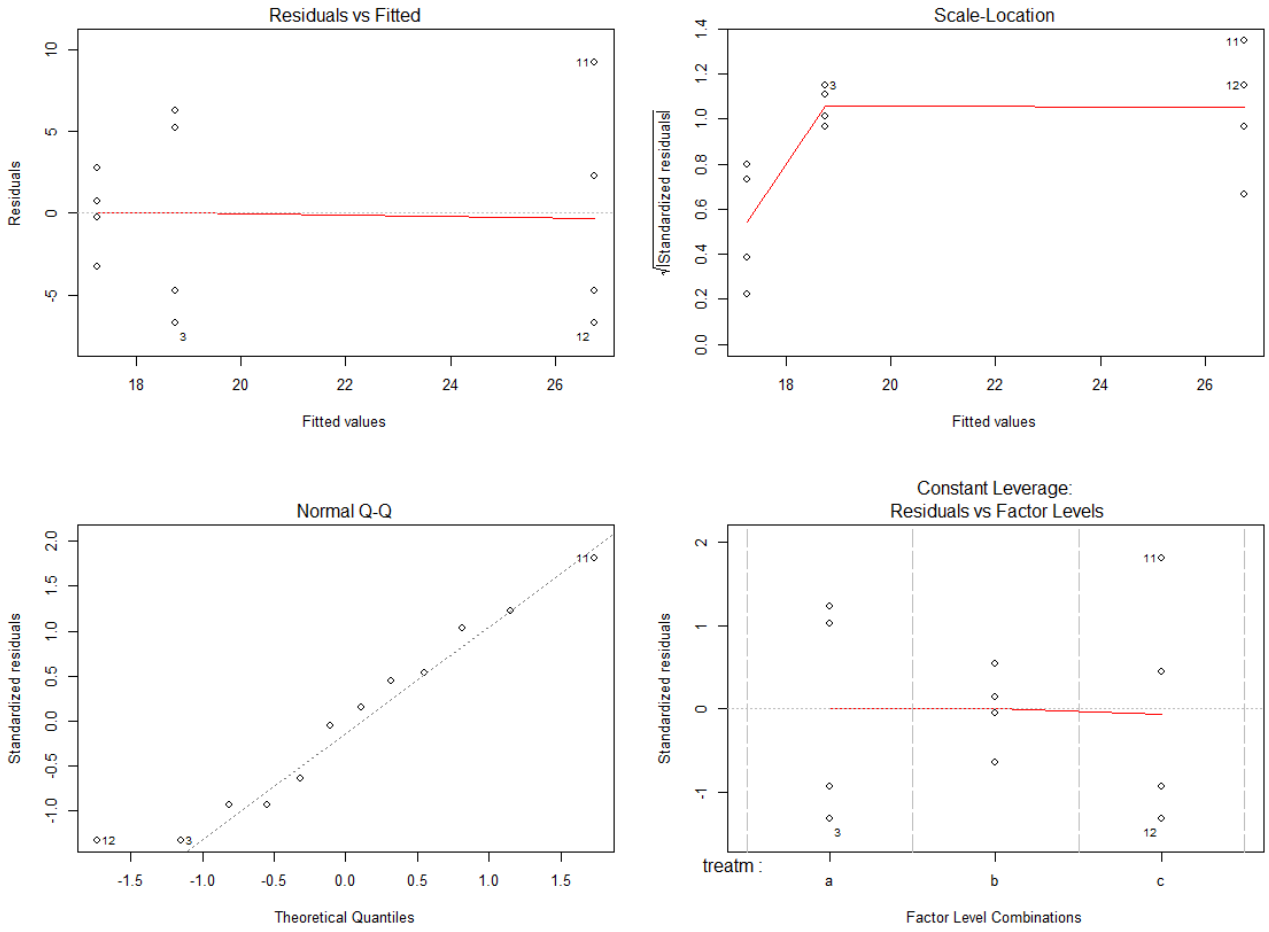
As shown above, the lm() function presented its findings a differently giving more information about significance of each treatment group. However, fundamentally the result was the same. The overall p-value of 0.09995 (for the F-statistic od 3.007 on 2 and 9 degrees of freedom) is higher than the significance level of 0.05. It means that we cannot reject the null hypothesis stating that the means for all three groups are equal, therefore treatments (A, B or C) did not have any effect on the resulting time to pain relief.

In order to assess validity of the model's results, I proceeded with the diagnostic and constructed the following four plots for both models (avm and avmodel) to visually assess normal distribution of the residuals and equal variance of the residuals.

```
> #diagnostic plots for avm constructed using aov() function
> layout(matrix(c(1,2,3,4),2,2)) # optional 4 graphs/page
> plot(avm)
```




```
> #diagnostic plots for avmodel constucted using lm() function
> layout(matrix(c(1,2,3,4),2,2)) # optional 4 graphs/page
> plot(avmodel)
```



As expected, the diagnostic graphs for both models are identical. In both cases they show that according to the Residuals vs. Fitted graph, the residuals variate around 0 and the red trend line is very close to horizontal. The second normal QQ graph checks for the normal distribution of the residuals, and it shows that they are very close to the normal QQ line. There are no real outliers that could have significantly distorted the models results.

Overall, we can conclude that the assumptions of normality, homoscedasticity and equal variance of the residuals are met and there are no compelling reasons not to trust the result of the ANOVA model.

Conclusions:

- ANOVA testing did not provide sufficient evidence allowing us to reject the null hypothesis stating that the means for all three treatment groups are equal. It means that treatments A, B, and C did not have any influence on the resulting time to pain relief.
- Since we did not reject the null hypothesis, we do not need to run any additional post-hoc tests.
- The ANOVA testing was conducted on a very small sample size. Each treatment group included only four data points which might be insufficient to reject the null hypothesis. In this case, the failure to

reject the null hypothesis does not necessarily mean the lack of treatments' influence on the dependent variable.