C2M3_Peer_Reviewed

February 13, 2022

1 C2M3: Peer Reviewed Assignment

1.0.1 Outline:

The objectives for this assignment:

1. Motivate the use of two-way ANOVA through real data analysis examples.
2. Interpret the two-way ANOVA model, with and without interaction terms.
3. Construct and interpret interaction plots to visually assess the importance of an interaction term.
4. Conduct hypothesis tests to decide whether a two-way ANOVA interaction term is statistically significant.
5. Use the two-way ANOVA and ANCOVA models to answer research questions using real data.

General tips:

1. Read the questions carefully to understand what is being asked.
2. This work will be reviewed by another human, so make sure that you are clear and concise in what your explanations and answers.

[36]:
```
# Load Required Packages
library(tidyverse)
library(ggplot2) # a package for nice plots!
library(dplyr)
library(emmeans)
```

2 Problem 1: e-reader data

In this assignment, we learn to answer our two-way ANOVA research questions through the analysis of real data. We will use the ereader data. The study that generated these data can be found here: P.-C. Chang, S.-Y. Chou, K.-K. Shieh (2013). "Reading Performance and Visual Fatigue When Using Electronic Displays in Long-Duration Reading Tasks Under Various Lighting Conditions," Displays, Vol. 34, pp. 208-214.)

Electronic paper display devices, such as the Amazon Kindle have changed the way that people read. But has it changed for the better? In a 2013 study titled "Reading Performance and Visual

Fatigue When Using Electronic Displays in Long-Duration Reading Tasks Under Various Lighting Conditions", researchers set out to ask whether reading speed (a continuous variable) differed across different electronic paper displays. In addition, they were also interested in whether different lighting conditions impacted reading speed. As such, this experiment had one response with two different factors:

1. Device type: three different types.
    1. Sony PRS-700 with a 6-in. display, 800×600 resolution;
    2. Amazon Kindle DX with a 9.7-in. display, 1200×824 resolution; and
    3. iRex 1000S with a 10.2-in. display, 1024×1280 resolution.
2. Lighting Condition: four different conditions (200Lx,500Lx, 1000Lx, 1500Lx), Lx = lux, one lumen per square meter
3. Reading Time: measured in seconds.

With these data, we might ask the following **research questions**:

1. Are the effects of device type significant? That is, is there evidence that suggests that individuals read at different speeds based on the device that they are using?
2. Are the effects of lighting conditions significant? That is, is there evidence that suggests that individuals read at different speeds based on the reading lighting conditions?
3. Do device type and lighting conditions *interact*? For example, Suppose that, on average, people can read for longer on device A than on device B, in low light. Is that trend the same in medium light, or bright light? If not, for example, if B is better than A in bright light, then type and lighting interact.

Through this entire analysis, let's set $\alpha = 0.05$.

First, let's read in the data, and store the appropriate variables as factors.

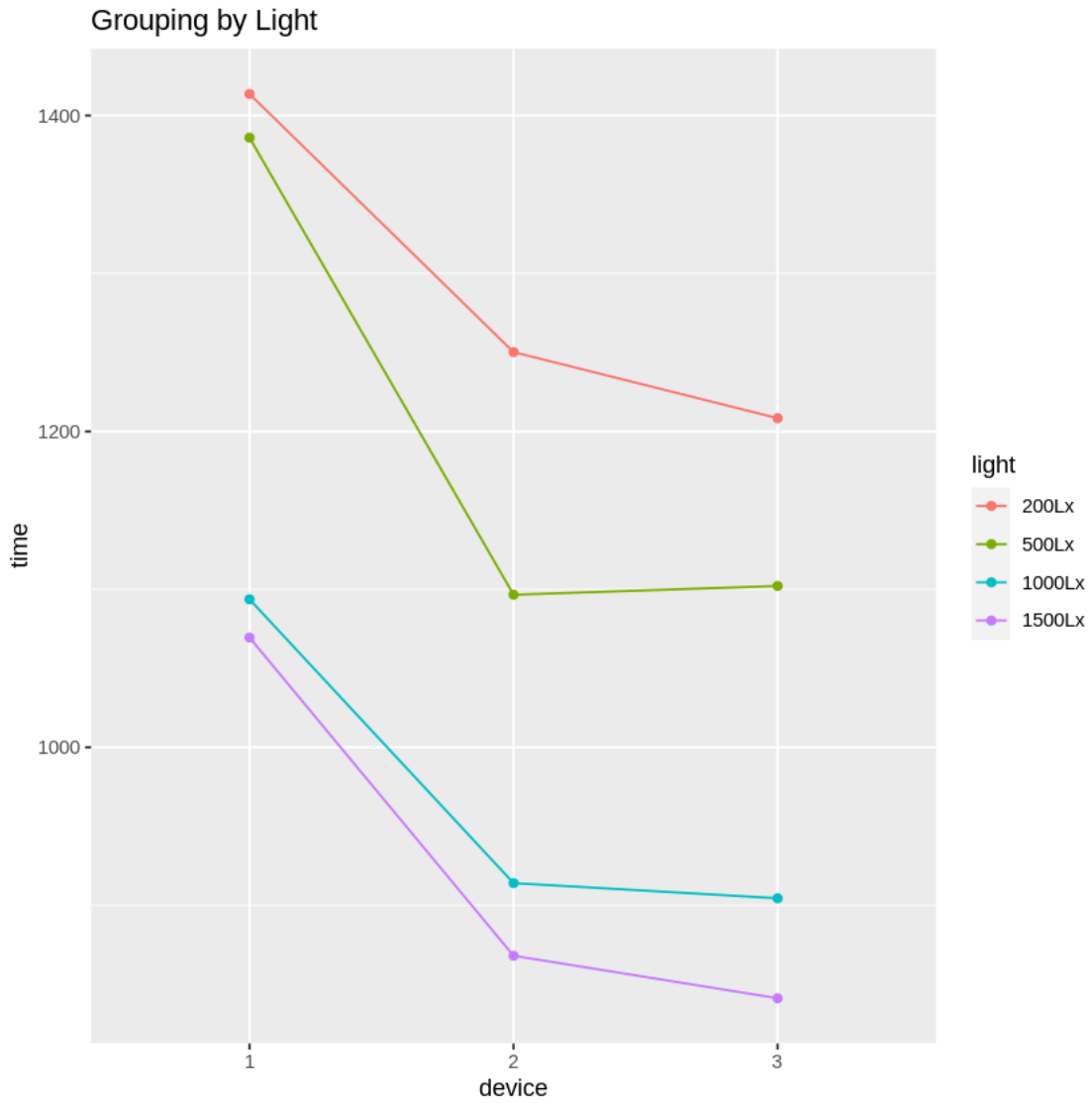
[37]:
```
# Load the data
read = read.csv("ereader.txt", sep="\t")

names(read) = c("device", "light", "time")
read$device = as_factor(read$device)
read$light = as.factor(read$light)
read$light = recode(read$light, "1" = "200Lx", "2" = "500Lx", "3" = "1000Lx",␣
↪"4" = "1500Lx")

summary(read)
```

```
 device    light         time
 1:19   200Lx :14   Min.   : 543.8
 2:20   500Lx :15   1st Qu.: 861.4
 3:20   1000Lx:15   Median :1105.4
        1500Lx:15   Mean   :1090.2
                    3rd Qu.:1300.0
                    Max.   :1797.2
```

1.(a) Construct interaction plots, and visually assess and comment on whether interactions are present.

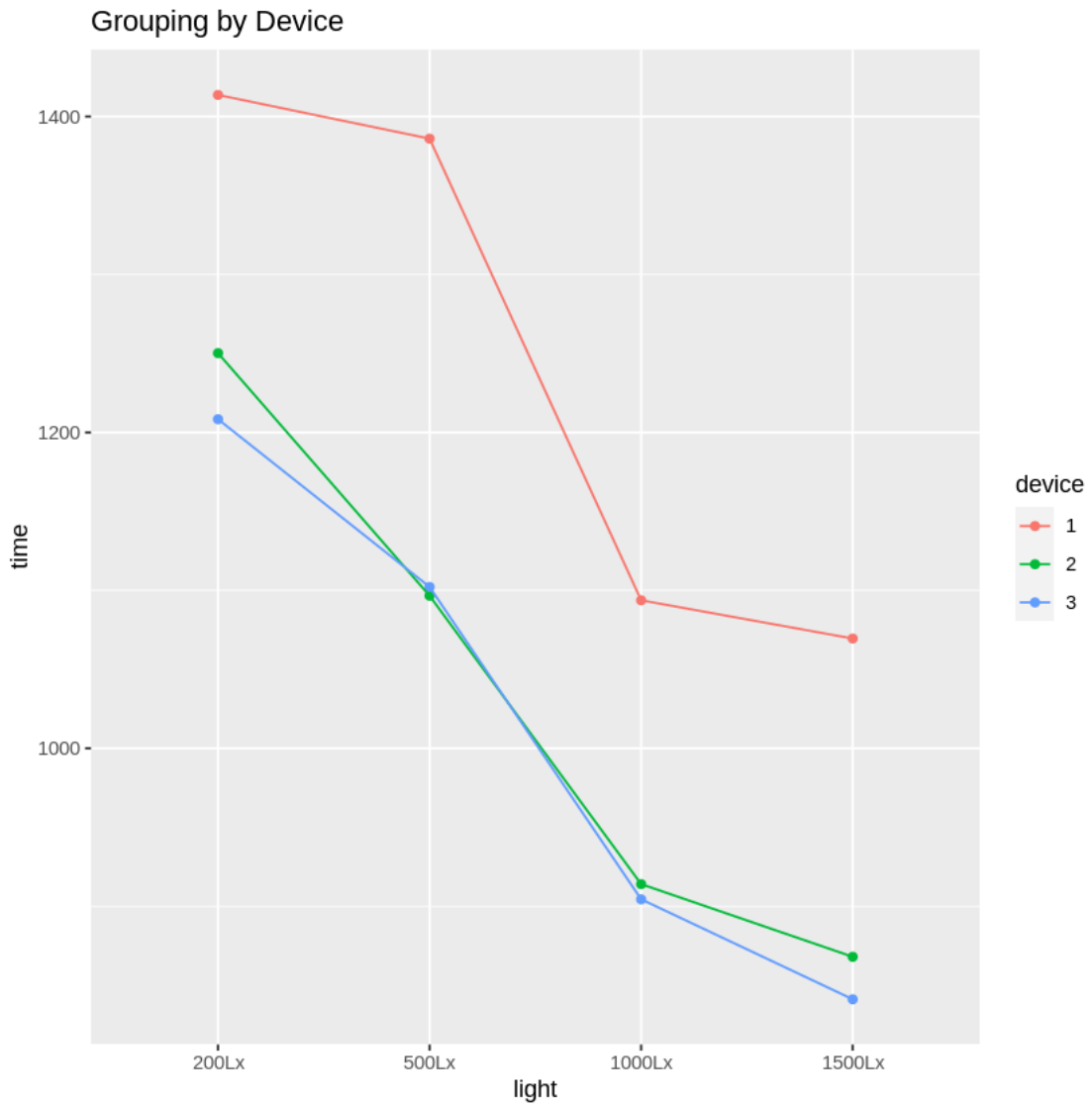
[38]:
```
# Your Code Here
p1 = ggplot(data=read, aes(x=device, y=time, group = light, color=light))
p1 = p1 + stat_summary(fun.y = mean, geom="point")
p1 = p1 + stat_summary(fun.y = mean, geom="line")
p1 + ggtitle('Grouping by Light')
```

```
Warning message:
"`fun.y` is deprecated. Use `fun` instead."
Warning message:
"`fun.y` is deprecated. Use `fun` instead."
```



[39]:
```
p2 = ggplot(data=read, aes(x=light, y=time, group = device, color=device))
p2 = p2 + stat_summary(fun.y = mean, geom="point")
p2 = p2 + stat_summary(fun.y = mean, geom="line")
p2 + ggtitle('Grouping by Device')
```

```
Warning message:
"`fun.y` is deprecated. Use `fun` instead."
Warning message:
"`fun.y` is deprecated. Use `fun` instead."
```



2.0.1 Answer:

Our first plot suggest no evidence of interaction as the lines are basically parallel to each other. However, the second plot states there may be very minimal interaction. For 200Lx lighint Device 2 has a longer reading time than device 3. However, at 500Lx they switch. Going to 1000Lx lighting, the comparison is similar to 200Lx where Device 2 has a longer reading time than device 3. Because these differences are minimal, it's hard to say if they are statistically significant.

1.(b) Now, let's formally test for an interaction. Fit a model with an interaction, and one without, and conduct an F-test. State the appropriate decision for the test.

[40]:
```
# Your Code Here
lm_noint = lm(time ~ light + device, data= read)
lm_int = lm(time ~ light + device + light:device, data= read)
summary(lm_noint)
print('Model summary of Model with Interaction Below---------- ')
summary(lm_int)
```

```
Call:
lm(formula = time ~ light + device, data = read)

Residuals:
   Min     1Q Median     3Q    Max
-500.0 -194.6  -24.8  204.9  460.5

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1438.25      87.22  16.489  < 2e-16 ***
light500Lx    -97.46      97.30  -1.002 0.321052
light1000Lx  -321.66      97.30  -3.306 0.001704 **
light1500Lx  -366.16      97.30  -3.763 0.000421 ***
device2      -209.73      83.89  -2.500 0.015547 *
device3      -227.93      83.89  -2.717 0.008879 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 261.7 on 53 degrees of freedom
Multiple R-squared:  0.3455,Adjusted R-squared:  0.2838
F-statistic: 5.596 on 5 and 53 DF,  p-value: 0.0003268



[1] "Model summary of Model with Interaction Below---------- "

Call:
lm(formula = time ~ light + device + light:device, data = read)

Residuals:
    Min      1Q  Median      3Q     Max
```

```
-497.41 -188.21  -17.28  207.16  463.53

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)          1413.63     138.44  10.211 1.62e-13 ***
light500Lx            -27.67     185.74  -0.149   0.8822
light1000Lx          -319.94     185.74  -1.723   0.0915 .
light1500Lx          -344.14     185.74  -1.853   0.0702 .
device2              -163.44     185.74  -0.880   0.3833
device3              -205.28     185.74  -1.105   0.2747
light500Lx:device2   -125.92     255.27  -0.493   0.6241
light1000Lx:device2   -16.24     255.27  -0.064   0.9495
light1500Lx:device2   -38.04     255.27  -0.149   0.8822
light500Lx:device3    -78.53     255.27  -0.308   0.7597
light1000Lx:device3    15.99     255.27   0.063   0.9503
light1500Lx:device3   -23.11     255.27  -0.091   0.9283
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 276.9 on 47 degrees of freedom
Multiple R-squared:  0.3502,Adjusted R-squared:  0.1981
F-statistic: 2.302 on 11 and 47 DF,  p-value: 0.02369
```

[41]:
```
anova(lm_noint, lm_int)
```

A anova: 2 × 6

	Res.Df <dbl>	RSS <dbl>	Df <dbl>	Sum of Sq <dbl>	F <dbl>	Pr(>F) <dbl>
1	53	3628970	NA	NA	NA	NA
2	47	3603108	6	25861.55	0.05622427	0.9992146

2.0.2 Answer:

With this test we are testing the following hypothesis: H_0: the reduced model is sufficient H_1: the reduced model is not sufficient In this case, we fail to reject the null hypothesis. The reduced model is sufficient, we do not need an interaction term.

1.(c) Before we interpret this model with respect to research question #1 above (just below the data description), let's decide whether the differences that the model reports are statistically significant.

Investigate this question using Bonferroni post hoc comparisons. That is, conduct all pairwise post hoc comparisons for device type using a Bonferroni correction and an overall type I error rate of $\alpha = 0.05$. Comment on the results.

[42]:
```
# Your Code Here
pairs(lsmeans(lm_noint, "device"), adjust="bonferroni")
```

```
 contrast estimate   SE df t.ratio p.value
 1 - 2       209.7 83.9 53 2.500   0.0466
 1 - 3       227.9 83.9 53 2.717   0.0266
 2 - 3        18.2 82.7 53 0.220   1.0000

Results are averaged over the levels of: light
P value adjustment: bonferroni method for 3 tests
```

2.0.3 Answer:

It seems that the device type 1 has a statistically significant difference in reading time when compared to device 2 and 3. However, the difference in device 2 and 3 are not statistically significant.

1.(d) Using the post hoc comparisons from above, let's focus on research question #1 from above: Are the effects of device type significant? That is, is there any evidence that suggests that individuals read faster or slower based on the device that they are using

2.0.4 Answer:

Looking at the original design of the experiment we have the device numbers to represent the following names: 1. Sony PRS-700 2. Amazon Kindle DX 3. iRex 100S

When comparing the Sony PRS-700 and Amazon Kindle DX, we can see the mean reading times are statistically significant; therefore, there is a difference in mean reading times (can read on Sony device 210 seconds longer than Amazon Kindle). When comparing the Sony PRS-700 and the iRex100S, we cann see the mean reading times are statistically significant; therefore, there is a difference in mean reading times (can read on Sony device 227.9 seconds longer than iRex device. When comparing the Amazon Kindle and the iRex100S, the mean reading times are not statistically significant; therefore, we cannot state there is a true difference in mean reading time between the two devices.

[ ]: