C2M3_Peer_Reviewed

July 1, 2023

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.

[1]:
```
# Load Required Packages
library(tidyverse)
library(ggplot2) # a package for nice plots!
library(dplyr)
library(emmeans)
```

```
  Attaching packages                                        tidyverse
1.3.0

  ggplot2 3.3.0       purrr   0.3.4
  tibble  3.0.1       dplyr   0.8.5
  tidyr   1.0.2       stringr 1.4.0
  readr   1.3.1       forcats 0.5.0

  Conflicts
tidyverse_conflicts()
  dplyr::filter() masks stats::filter()
```

```
dplyr::lag()    masks stats::lag()
```

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.

[2]:
```
# 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
```

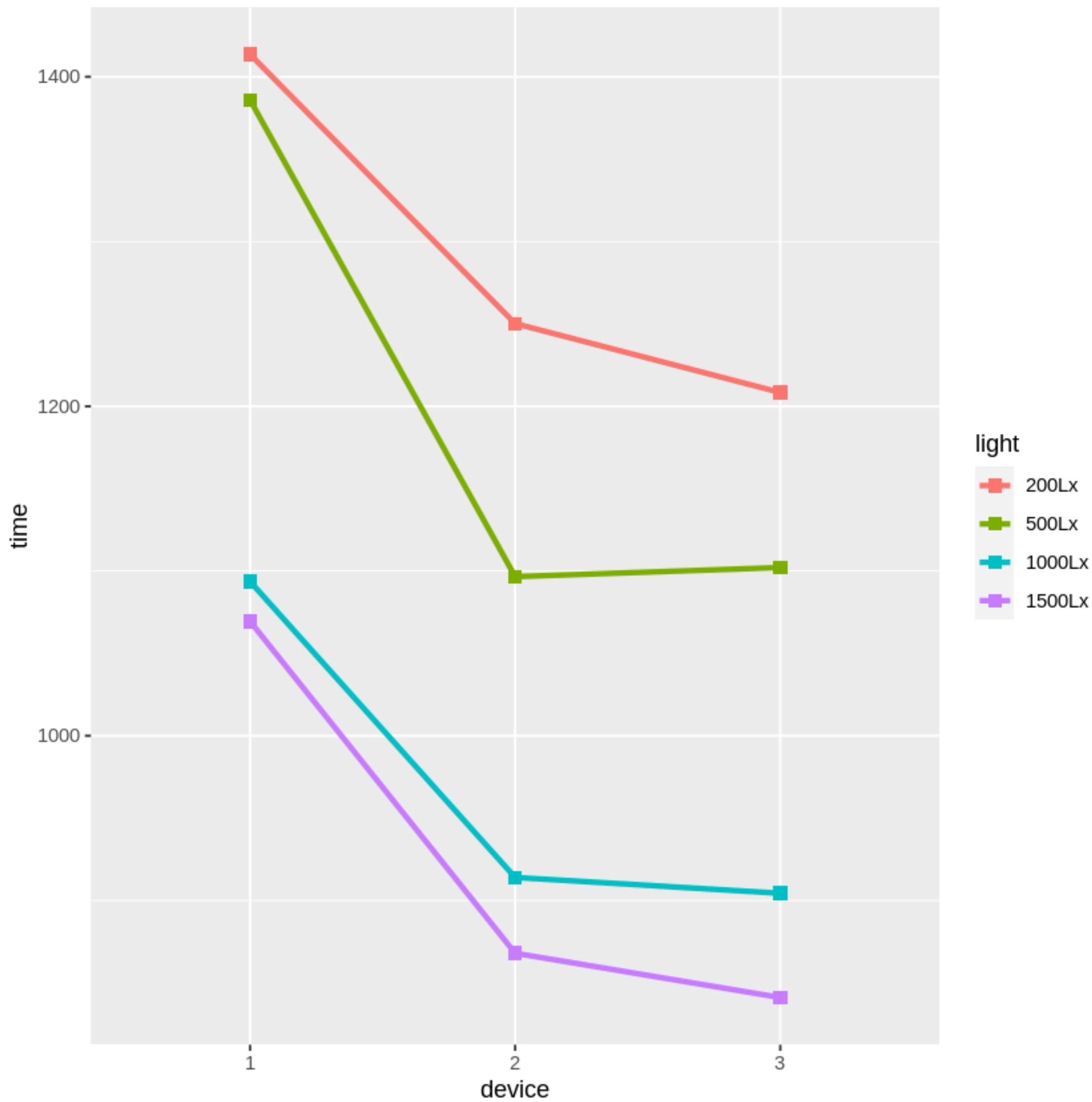
1.(a) Construct interaction plots, and visually assess and comment on whether interactions are present.

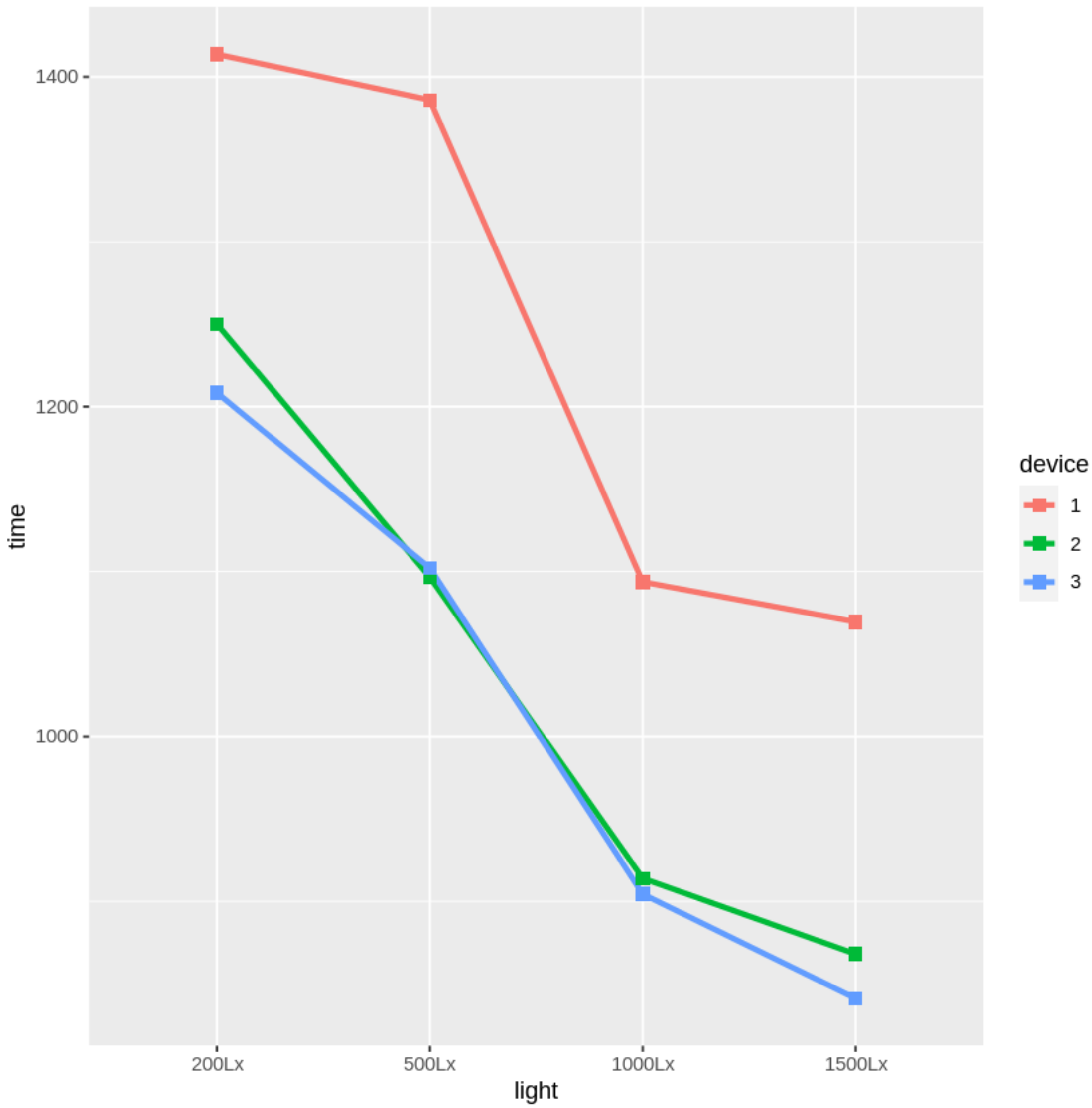
[3]:
```
# Your Code Here

mean_time = read %>% group_by(device, light) %>% summarise(time=mean(time))

ggplot(data=mean_time, aes(x=device, y=time)) +
    geom_line(size = 1.2, aes(group=light, color=light)) +
    geom_point(size = 2.6, aes(color=light), shape=15)

ggplot(data=mean_time, aes(x=light, y=time)) +
    geom_line(size = 1.2, aes(group=device, color=device)) +
    geom_point(size = 2.6, aes(color=device), shape=15)
```


1400
1200
1000
time
1
2
3
device
light
200Lx
500Lx
1000Lx
1500Lx



There are some possible ***weak*** indications of interactions as we see large deviations in slope for 500Lx when moving from device 1 to 2 on the first plot. Similarly for device 1 the slope is slightly different when moving from 200Lx to 500Lx. But that's very likely due to the sampling errors.

Also, it looks like device 1 ranks consistently higher in terms of reading times when compared to devices 2 and 3 across all lighting conditions.

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.

[4]:
```
# Your Code Here
m = lm(time ~ device + light, data=read)
m_int = lm(time ~ device + light + device * light, data=read)
```

```r
summary(m)
summary(m_int)

anova(m, m_int)
```

```
Call:
lm(formula = time ~ device + light, 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 ***
device2      -209.73      83.89  -2.500 0.015547 *  
device3      -227.93      83.89  -2.717 0.008879 ** 
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 ***
---
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



Call:
lm(formula = time ~ device + light + device * light, 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 ***
device2              -163.44     185.74  -0.880   0.3833    
device3              -205.28     185.74  -1.105   0.2747    
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:light500Lx   -125.92     255.27  -0.493   0.6241    
device3:light500Lx    -78.53     255.27  -0.308   0.7597    
device2:light1000Lx   -16.24     255.27  -0.064   0.9495    
device3:light1000Lx    15.99     255.27   0.063   0.9503    
```

```
device2:light1500Lx   -38.04     255.27  -0.149   0.8822
device3:light1500Lx   -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
```

A anova: 2×6

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

We can't reject the null (reduced model is sufficient) as the p-value $= 0.999 > 0.05$ therefore the interactions are not statistically significant.

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.

[5]:
```
# Your Code Here
pairs(lsmeans(m, '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
```

We see that the data suggest significant (at 5% level) differences between devices 1 and 2 and devices 1 and 3 as both p-values 0.0466 and 0.0266 are less than 0.05. This result confirms visual findings from the interaction plots.

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

It looks like there is some basis for that claim. Indeed, individuals tend to read about 210 seconds (3.5 minutes) faster when using devices 2 and 3 as opposed to using device 1.

[6]:
```
# Your Code Here
```