C2M3: Peer Reviewed Assignment

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.

In [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()
```

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.])

Processing math: 100%

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.
   A. Sony PRS-700 with a 6-in. display, 800 × 600 resolution;
   B. Amazon Kindle DX with a 9.7-in. display, 1200 × 824 resolution; and
   C. iRex 1000S with a 10.2-in. display, 1024 × 1280 resolution.

1. Lighting Condition: four different conditions (200Lx,500Lx, 1000Lx, 1500Lx), Lx = lux, one lumen per square meter

1. 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?

1. 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?

1. 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 α = 0.05.

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

In [28]:
```
# 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" = "1000L
x", "4" = "1500Lx")

summary(read)
head(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  
```

A data.frame: 6 × 3

	device	light	time
	<fct>	<fct>	<dbl>
1	1	200Lx	1405.92
2	1	200Lx	1797.21
3	1	200Lx	1155.96
4	1	200Lx	1295.44
5	1	500Lx	1022.32
6	1	500Lx	1538.07

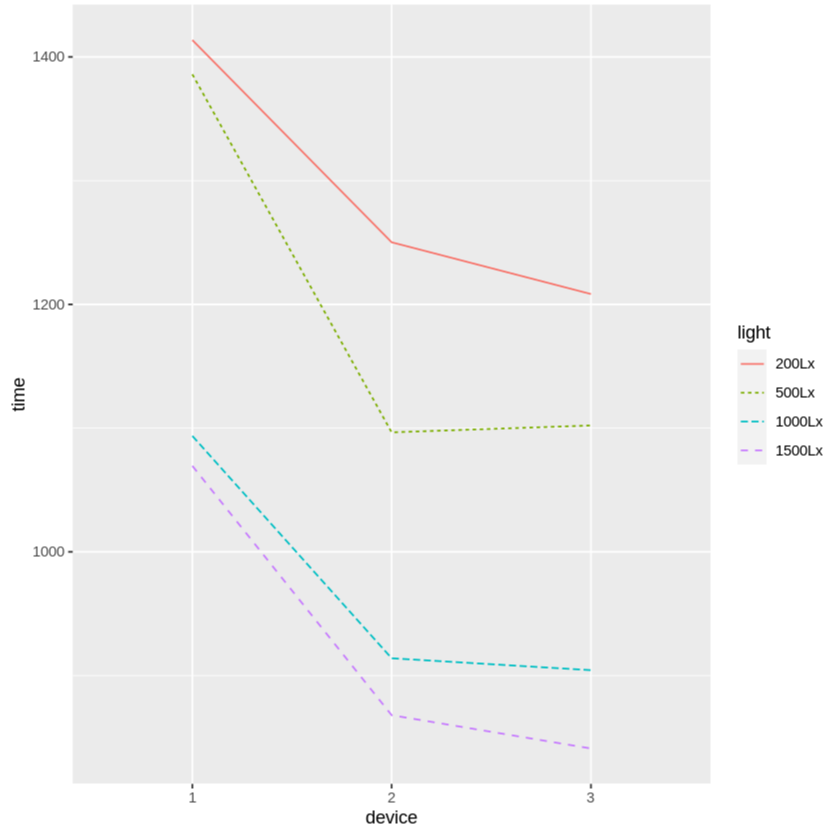
1.(a) Construct interaction plots, and visually assess and comment on whether interactions are present.

In [34]:
```
# Your Code Here

summ = aggregate(time ~ light + device, data = read, FUN = mean)
summ
p1 = ggplot(data = summ, aes(x = device, y = time, group = light, linetype
 = light)) + geom_line(aes(color = light))
p1
```

A data.frame: 12 × 3

light	device	time
<fct>	<fct>	<dbl>
200Lx	1	1413.633
500Lx	1	1385.960
1000Lx	1	1093.694
1500Lx	1	1069.490
200Lx	2	1250.188
500Lx	2	1096.594
1000Lx	2	914.006
1500Lx	2	868.010
200Lx	3	1208.350
500Lx	3	1102.146
1000Lx	3	904.400
1500Lx	3	841.100



Solution

The plot shows a shorter mean time as we move from device 1 to 3 for all the light settings except for 500 Lx. There is not enough information to determine if there are any interactions.

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.

In [39]:

```
# Your Code Here
lmod1 = lm(time ~ device + light, data = read)
summary(lmod1)

lmod2 = lm(time ~ device + light + device:light, data = read)
summary(lmod2)

anova(lmod1, lmod2)
```

```
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

Solution
The table produced from the two way ANOVA shows that the p-value is 0.99, much larger than the assumed alpha level of 0.05. This means that we fail to reject the null hypothesis and conclude that there is no interaction. Also, the table summarizing the linear model without interactions confirms that a model without the interaction terms is sufficient, having a p-value of 0.003.

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 α = 0.05. Comment on the results.

In [42]:
```
# Your Code Here
pairs(lsmeans(lmod1, "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
```

Solution The two pairs that are statistically significant are: 1-2 and 1-3, having p-values of 0.0466 and 0.0266, respectively.

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

Solution
The research question of concern is:

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?

Based on the analyses above, we can see that there is a statistically significant difference in terms of reading time between devices 1 and 2 (PRS-700 and Amazon Kindle). The estimated difference is about 210 seconds, while having a low p-value of 0.0466, slightly below the required alpha-level. Similarly, the difference in reading time between devices 1 and 3 (PRS-700 and iRex) is also statistically significant, with a difference of about 228 seconds and a low p-value of 0.0266.
The difference between the Kindle and iRex is not significant, since not only is the difference itself small (18 seconds), but also with a very large p-value.

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