Lesson 9:
Analysis of variance: Two-Way ANOVA

Victoria Blanes-Vidal

Manuella Lech Cantuaria

The Maersk Mc-Kinney Møller Institute

Applied AI and Data Science

Lesson	Week	Date	TOPICS	Teacher
1	35	1/Sep	Introduction to the course Descriptive statistics –Part I	MLC
2	36	8/sep	Descriptive statistics –Part II	MLC
3	37	15/Sep	Probability distributions	MLC
4	38	22/Sep	Hypothesis testing (one sample)	VBV
5	39	29/Sep	Hypothesis testing (two samples)	VBV
6	40	6/Oct	ANOVA one-way	VBV
7	41	13/Oct	R class (Introduction to R and descriptive statistics) Point giving activity	MLC
-	42	20/Oct	NO CLASS (Autum holidays)	
8	43	27/Oct	R class (hypothesis testing + ANOVA)	MLC
9	44	3/Nov	ANOVA two-way	VBV
-	45	10/Nov	NO CLASS	
10	46	17/Nov	Regression analysis	VBV
11	47	24/Nov	Multiple regression Point giving activity	MLC
12	48	1/Dec	Notions of experimental design and questions	VBV+MLC

VBV = Victoria Blanes-Vidal
MLC = Manuella Lech Cantuaria

Analysis of variance: Two-way ANOVA

1. Recap: One-way Analysis of variance
2. What is “Two-Way Analysis of Variance”?
3. How can I interpret the results of a “Two-way ANOVA”?
    1. Example Magazines (Dataset 1)
    2. Example Magazines (Dataset 2)
    3. Example Magazines (Dataset 3)
    4. Example Magazines (Dataset 4)
4. In-class exercise

Analysis of variance: Two-way ANOVA

1. **Recap: One-way Analysis of variance**
2. What is “Two-Way Analysis of Variance”?
3. How I interpret the results of a “Two-way ANOVA”?
    1. Example Magazines (Dataset 1)
    2. Example Magazines (Dataset 2)
    3. Example Magazines (Dataset 3)
    4. Example Magazines (Dataset 4)
4. In-class exercise

The <u>one-way</u> analysis of variance is used when we want to study the effect of **one** factor (or "treatment") on a response variable.

For example… we want to study whether the cover of a magazine has an effect on the sales of the magazine:

Since we want to study the effect of only 1 Factor (Cover) on the response variable (Sales)…

We will then use a <u>One-way ANOVA</u>

In a One-way ANOVA we get a Table like this one…

```
> One_way_anova_magazine = aov(Sales ~ Cover,
data=Sales_Cover)
> summary(One_way_anova_magazine)
             Df Sum Sq Mean Sq F value Pr(>F)
Cover         2  57.33   28.67   2.455   0.12
Residuals    15 175.17   11.68
```

If the p-value is > level of significance (e.g. 0.05) *then:*
There is NOT a significant effect of the Cover on the sales (that is, the sales DO NOT differ depending on the cover)

```
> One_way_anova_magazine = aov(Sales ~ Cover,
data=Sales_Cover)
> summary(One_way_anova_magazine)
             Df Sum Sq Mean Sq F value  Pr(>F)
Cover         2  256.4  128.22   10.39 0.00148 **
Residuals    15  185.2   12.34
```

If the p-value is < level of significance (e.g. 0.05) *then:*
There is a significant effect of the Cover on the sales (that is, the sales differ depending on the cover)

Analysis of variance: Two-way ANOVA

1. **Recap: One-way Analysis of variance**
2. **What is “Two-Way Analysis of Variance”?**
3. How can I interpret the results of a “Two-way ANOVA”?
    1. Example 1
    2. Example 2
    3. Example 3
    4. Example 4
4. In-class exercise

The **two-way ANOVA** is an extension of the one-way analysis of variance; it allows to study the effect of <u>two factors</u> (or treatments).

Let's imagine that, apart from studying the effect of the magazine cover on sales,

Factor A:
Cover layout

Cover 1

Cover 2

Cover 3

<u>we also want to study</u> whether the location in which we display the magagine, has an effect on the sales of the magazine:

Factor B:
Location

Location 1: Low

Location 2: High

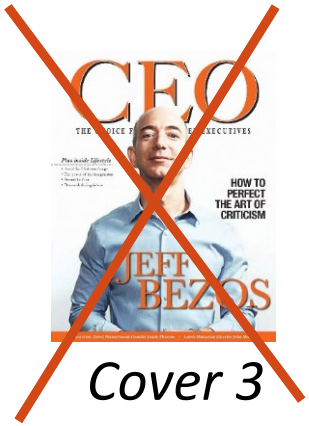
Since we want to study the effect of <u>2 Factors</u> (Cover layout and Location) on the response variable (Sales)…

We will then use a <u>Two-way ANOVA</u>

Factor A:
Cover layout

Cover 1

Cover 2



Cover 3

Factor B:
Location

Location 1: Low

Location 2: High

In this course we will only learn

Two-way ANOVA, when each of the two factors has ***two levels***

(not more than 2 levels)

The **two-way ANOVA** allows us to answer 3 questions:

1. Does the cover have an effect on the sales?

Factor A: Cover layout

Cover 1

Cover 2

2. Does the location have an effect on the sales?

Factor B: Location

Location 1: Low

Location 2: High

3. Is there any interaction between the effect of the cover and the location, on the sales?

Factor A x B: Interaction

The **two-way ANOVA** has 3 null hypotheses (and 3 alternative hypotheses):

Null hypothesis: The cover does not have an effect on sales (sales do not differ depending on the cover)

Alternative hypothesis: The cover has effect on sales (sales do differ depending on the cover)

Null hypothesis: The location has no effect on sales (sales do not differ depending on the location)

Alternative hypothesis: The location has an effect on sales (sales do differ depending on the location)

Null hypothesis: There is not an interaction effect.

Alternative hypothesis: There is an interaction effect.

Analysis of variance: Two-way ANOVA

1. **Recap: One-way Analysis of variance**

2. **What is “Two-Way Analysis of Variance”?**

3. **How can I interpret the results of a “Two-way ANOVA”?**
    1. Example Magazines (Dataset 1)
    2. Example Magazines (Dataset 2)
    3. Example Magazines (Dataset 3)
    4. Example Magazines (Dataset 4)

4. In-class exercise

Analysis of variance: Two-way ANOVA

1. **Recap: One-way Analysis of variance**
2. **What is “Two-Way Analysis of Variance”?**
3. **How can I interpret the results of a “Two-way ANOVA”?**
    1. **Example Magazines (Dataset 1)**
    2. Example Magazines (Dataset 2)
    3. Example Magazines (Dataset 3)
    4. Example Magazines (Dataset 4)
4. In-class exercise

Two-way analysis of variance : Example

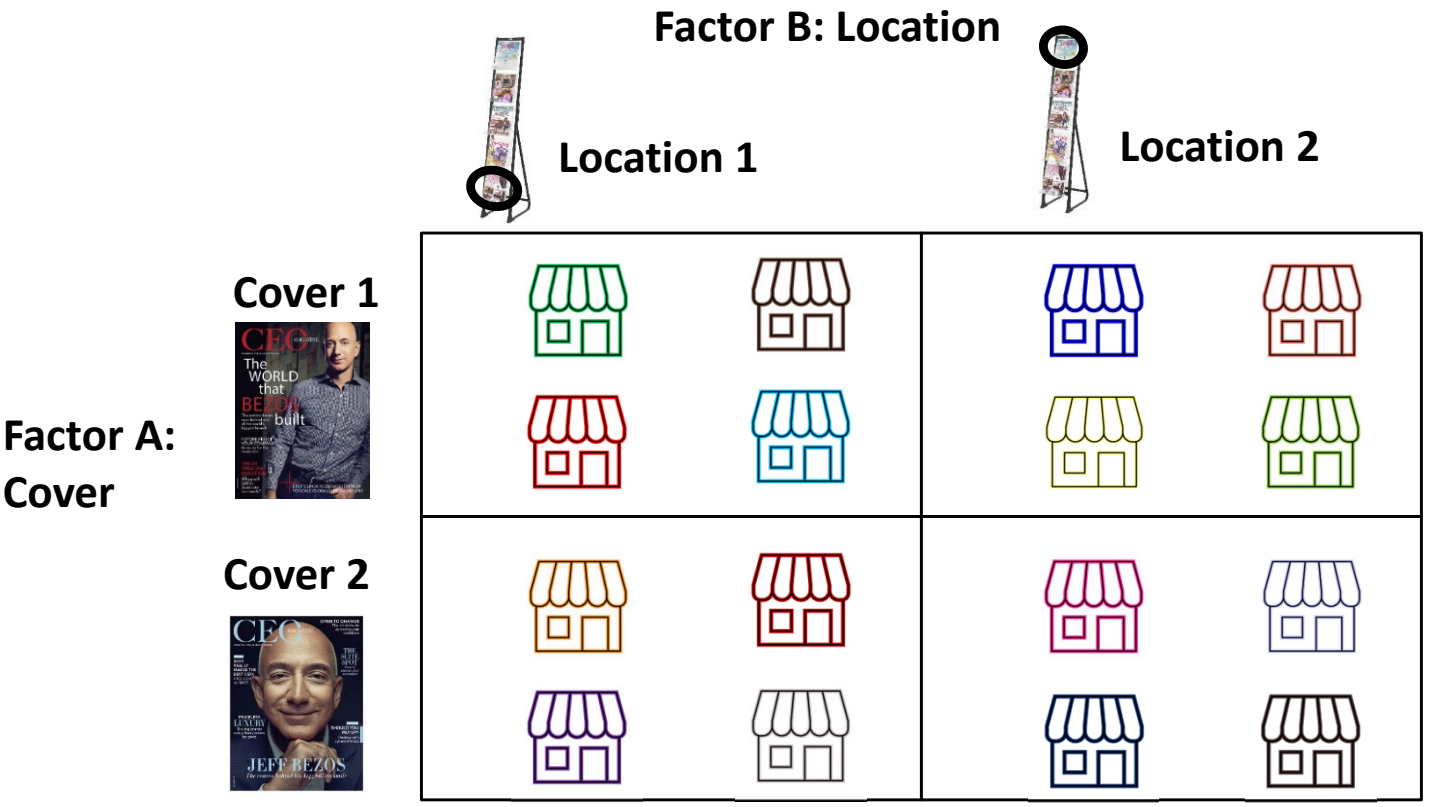
We want to study whether:

- the cover of the magazine (Cover 1 vs. Cover 2)
- the location of the magazine (Location 1 vs. Location 2)

⇨ Has an effect on the sales of the magazine.

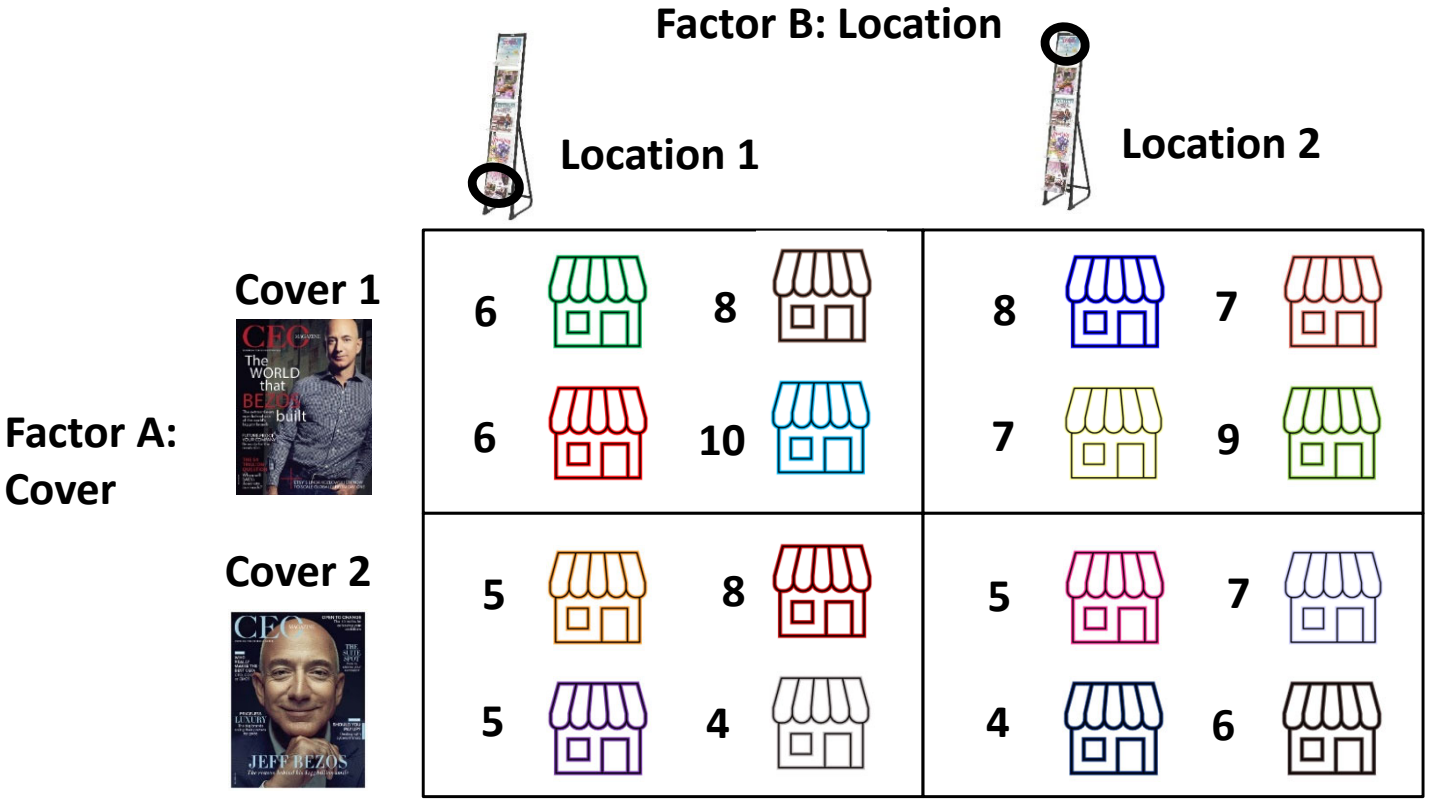
We design the following experiment:

- Four shops will sell Cover 1, placed in Location 1
- Four shops will sell Cover 1, placed in Location 2
- Four shops will sell Cover 2, placed in Location 1
- Four shops will sell Cover 2, placed in Location 2



Two-way analysis of variance : Example 1

After two weeks, we collect data on how many magazines were sold at each shop:



Does the cover have any effect on sales? Which cover is better?

Is there any effect of the Location of the magazine, on sales? Where should we display the magazine? At the top? At the bottom?

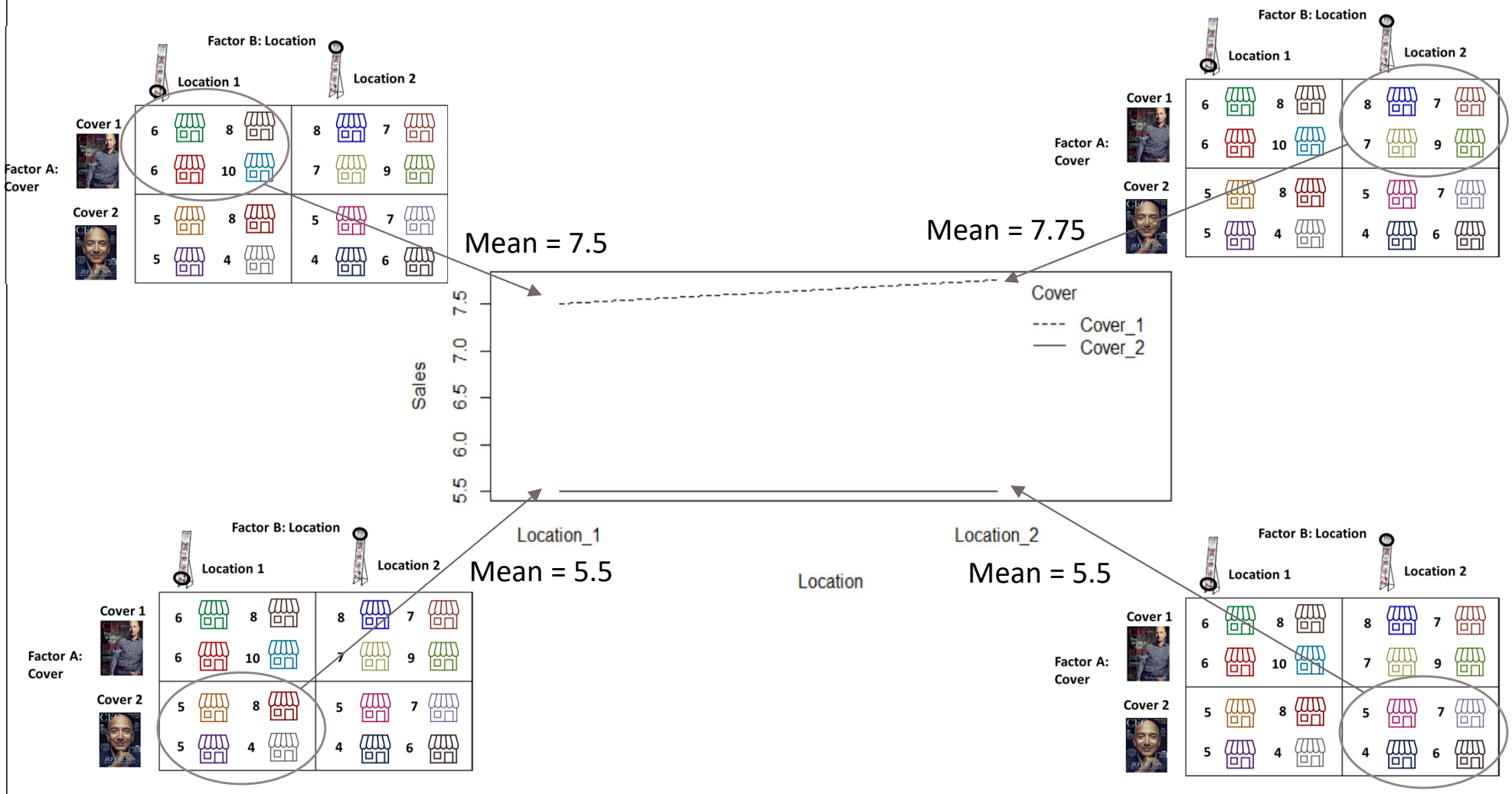
All in all, what would be your recommendation to the editor?

```
> Magazines_1 <-
read.table("C:/Users/vbv/Desktop/My_documents/Teaching/Teaching/
Statistical_Data_Analysis/2022_2023/Exercises_in_R/Magazines_Exa
mple_1.txt", header=TRUE)
> Magazines_1
     Cover   Location Sales
1  Cover_1 Location_1     6
2  Cover_1 Location_1     8
3  Cover_1 Location_1     6
4  Cover_1 Location_1    10
5  Cover_1 Location_2     8
6  Cover_1 Location_2     7
7  Cover_1 Location_2     7
8  Cover_1 Location_2     9
9  Cover_2 Location_1     5
10 Cover_2 Location_1     8
11 Cover_2 Location_1     5
12 Cover_2 Location_1     4
13 Cover_2 Location_2     5
14 Cover_2 Location_2     7
15 Cover_2 Location_2     4
16 Cover_2 Location_2     6
```

```
> Two_way_anova_Magazines_1 = aov(Sales ~ Cover + Location + Cover*Location,
data=Magazines_1)
> summary(Two_way_anova_Magazines_1)
               Df Sum Sq Mean Sq F value Pr(>F)
Cover           1 18.063  18.063   7.811 0.0162 *
Location        1  0.062   0.062   0.027 0.8722
Cover:Location  1  0.062   0.062   0.027 0.8722
Residuals      12 27.750   2.312
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

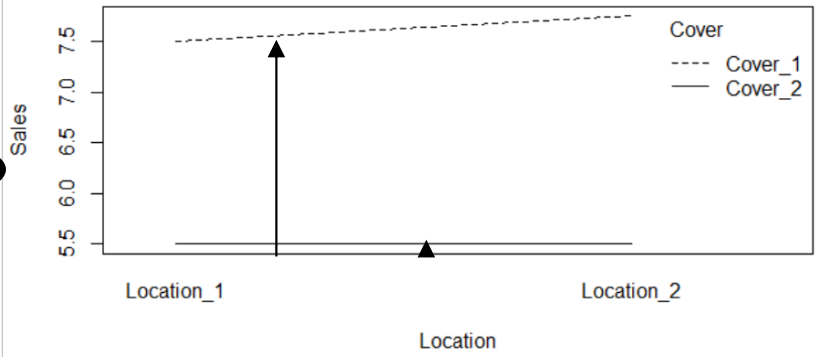
In order to better interpret the results in a Two-way ANOVA, we should plot the "interaction plot":

```
> interaction.plot(Magazines_1$Location, Magazines_1$Cover,
Magazines_1$Sales, xlab = "Location", ylab = "Sales", trace.label="Cover")
```



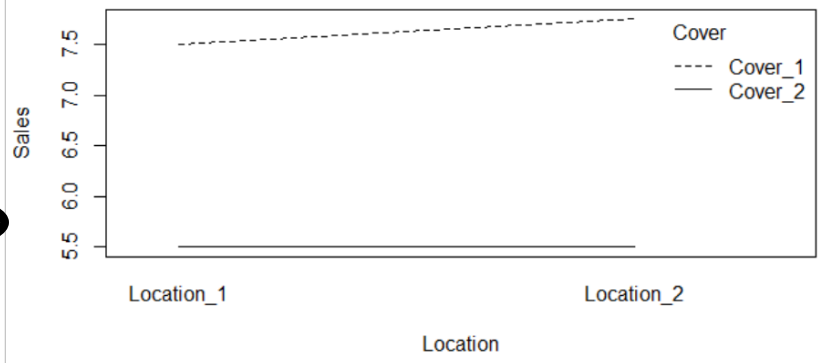
This plot displays means for the levels of one factor on the x-axis and a separate line for each level of another factor.

```
> Two_way_anova_Magazines_1 = aov(Sales ~ Cover + Location + Cover*Location, data=Magazines_1)
> summary(Two_way_anova_Magazines_1)
                Df Sum Sq Mean Sq F value Pr(>F)
Cover            1 18.063  18.063   7.811 0.0162 *
Location         1  0.062   0.062   0.027 0.8722
Cover:Location   1  0.062   0.062   0.027 0.8722
Residuals       12 27.750   2.312
```



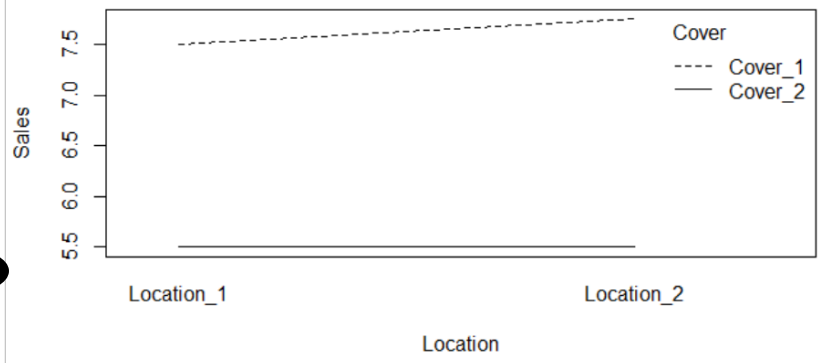
Effect of…	Conclusions from ANOVA	Conclusions from interaction plot
Cover	*Since p-value<0.05:* *"The magazine cover has a significant effect on sales"*	*"Sales are significantly higher when the cover is Cover 1, compared to when the Cover is Cover 2"*
Location		
Interaction		

```
> Two_way_anova_Magazines_1 = aov(Sales ~ Cover +
Location + Cover*Location, data=Magazines_1)
> summary(Two_way_anova_Magazines_1)
                Df Sum Sq Mean Sq F value Pr(>F)
Cover            1 18.063  18.063   7.811 0.0162 *
Location         1  0.062   0.062   0.027 0.8722
Cover:Location   1  0.062   0.062   0.027 0.8722
Residuals       12 27.750   2.312
```



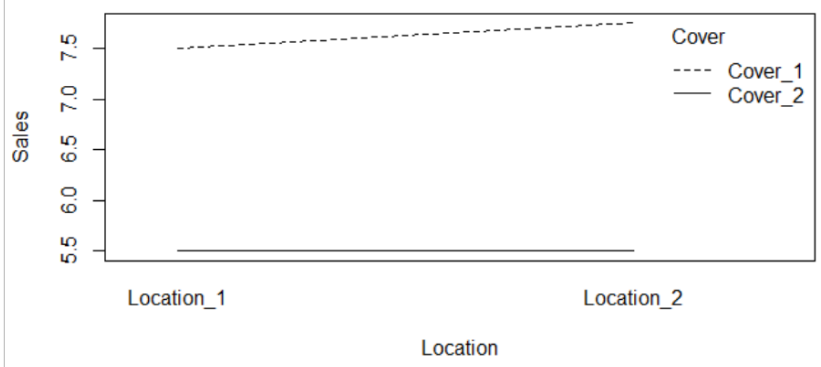
Effect of…	Conclusions from ANOVA	Conclusions from interaction plot
Cover	*Since p-value<0.05:* *"The magazine cover has a significant effect on sales"*	*"Sales are significantly higher when the cover is Cover 1, compared to when the Cover is Cover 2"*
Location	*Since p-value > 0.05:* *"The location does not have a significant effect on sales"*	-
Interaction		

```
> Two_way_anova_Magazines_1 = aov(Sales ~ Cover + Location + Cover*Location, data=Magazines_1)
> summary(Two_way_anova_Magazines_1)
                Df Sum Sq Mean Sq F value Pr(>F)
Cover            1 18.063  18.063   7.811 0.0162 *
Location         1  0.062   0.062   0.027 0.8722
Cover:Location   1  0.062   0.062   0.027 0.8722
Residuals       12 27.750   2.312
```



Effect of…	Conclusions from ANOVA	Conclusions from interaction plot
Cover	*Since p-value<0.05: "The magazine cover has a significant effect on sales"*	*"Sales are significantly higher when the cover is Cover 1, compared to when the Cover is Cover 2"*
Location	*Since p-value > 0.05: "The location does not have a significant effect on sales"*	-
Interaction	*Since p-value > 0.05: "The interaction effect is not significant"*	-

```
> Two_way_anova_Magazines_1 = aov(Sales ~ Cover + Location + Cover*Location, data=Magazines_1)
> summary(Two_way_anova_Magazines_1)
                Df Sum Sq Mean Sq F value Pr(>F)
Cover            1 18.063  18.063   7.811 0.0162 *
Location         1  0.062   0.062   0.027 0.8722
Cover:Location   1  0.062   0.062   0.027 0.8722
Residuals       12 27.750   2.312
```



Effect of…	Conclusions from ANOVA	Conclusions from interaction plot
Cover	*Since p-value<0.05:* *"The magazine cover has a significant effect on sales"*	*"Sales are significantly higher when the cover is Cover 1, compared to when the Cover is Cover 2"*
Location	*Since p-value > 0.05:* *"The location does not have a significant effect on sales"*	-
Interaction	*Since p-value > 0.05:* *"The interaction effect is not significant"*	-

My recommendation to the magazine editor:

If you want to sell more magazines:

Use Cover 1.

It does not matter whether the magazine is placed in Location 1 (low) or in Location 2 (high)

Let's practice with another dataset…

Dataset_2

Factor B: Location

Location 1

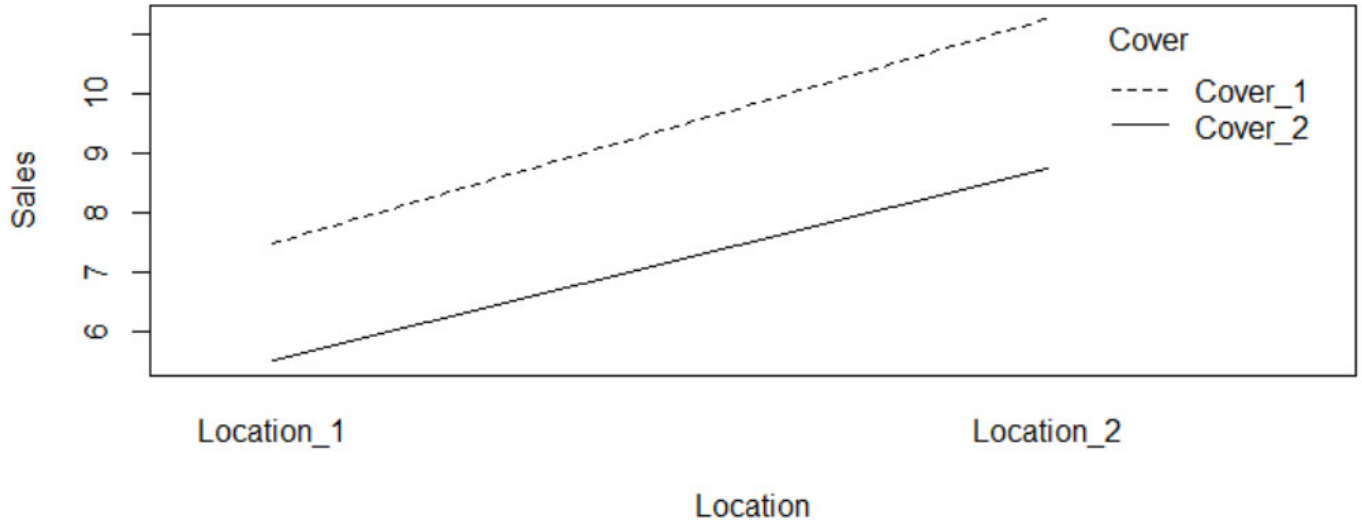
Location 2

Factor A: Cover	**Location 1**	**Location 2**
Cover 1 	6, 8, 6, 10	11, 9, 12, 13
Cover 2 	5, 8, 5, 4	7, 10, 10, 8

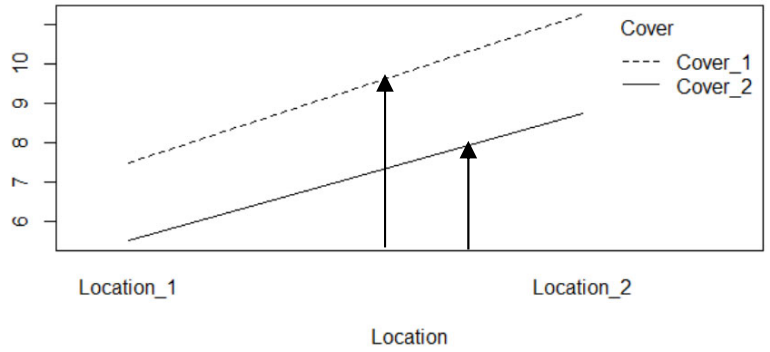
```
> Magazines_2 <-
read.table("C:/Users/vbv/Desktop/My_documents/Teaching/Teaching/S
tatistical_Data_Analysis/2022_2023/Exercises_in_R/Magazines_Examp
le_2.txt", header=TRUE)
> Magazines_2
     Cover   Location Sales
1  Cover_1 Location_1     6
2  Cover_1 Location_1     8
3  Cover_1 Location_1     6
4  Cover_1 Location_1    10
5  Cover_1 Location_2    11
6  Cover_1 Location_2     9
7  Cover_1 Location_2    12
8  Cover_1 Location_2    13
9  Cover_2 Location_1     5
10 Cover_2 Location_1     8
11 Cover_2 Location_1     5
12 Cover_2 Location_1     4
13 Cover_2 Location_2     7
14 Cover_2 Location_2    10
15 Cover_2 Location_2    10
16 Cover_2 Location_2     8
```

```
> Two_way_anova_Magazines_2 = aov(Sales ~ Cover + Location + Cover*Location,
data=Magazines_2)
> summary(Two_way_anova_Magazines_2)
               Df Sum Sq Mean Sq F value  Pr(>F)
Cover           1  20.25   20.25   6.845 0.02254 *
Location        1  49.00   49.00  16.563 0.00155 **
Cover:Location  1   0.25    0.25   0.085 0.77624
Residuals      12  35.50    2.96
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
> interaction.plot(Magazines$Location, Magazines$Cover, Magazines$Sales,
xlab = "Location", ylab = "Sales", trace.label="Cover")
```

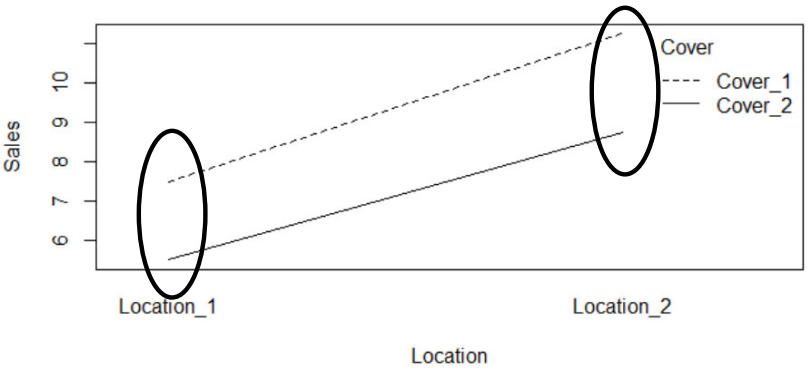


```
> Two_way_anova_Magazines_2 = aov(Sales ~ Cover +
Location + Cover*Location, data=Magazines_2)
> summary(Two_way_anova_Magazines_2)
                Df Sum Sq Mean Sq F value  Pr(>F)
Cover            1  20.25   20.25   6.845 0.02254 *
Location         1  49.00   49.00  16.563 0.00155 **
Cover:Location   1   0.25    0.25   0.085 0.77624
Residuals       12  35.50    2.96
```



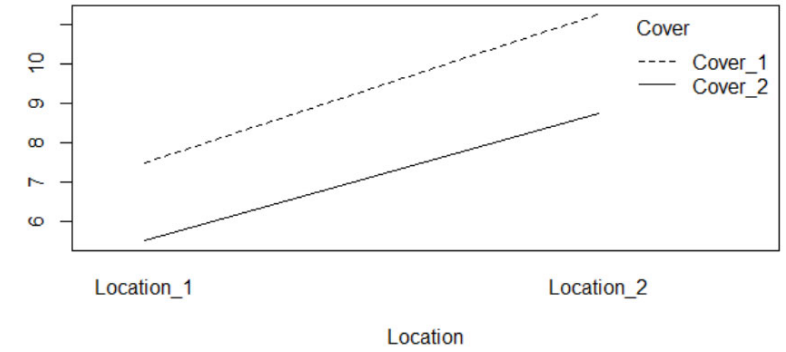
Effect of…	Conclusions from ANOVA	Conclusions from interaction plot
Cover	*Since p-value < 0.05 : ”The magazine cover has a significant effect on sales”*	*”Sales are significantly higher when the cover is Cover 1, compared to when the Cover is Cover 2”*
Location		
Interaction		

```
> Two_way_anova_Magazines_2 = aov(Sales ~ Cover +
Location + Cover*Location, data=Magazines_2)
> summary(Two_way_anova_Magazines_2)
                Df Sum Sq Mean Sq F value  Pr(>F)
Cover            1  20.25   20.25   6.845 0.02254 *
Location         1  49.00   49.00  16.563 0.00155 **
Cover:Location   1   0.25    0.25   0.085 0.77624
Residuals       12  35.50    2.96
```



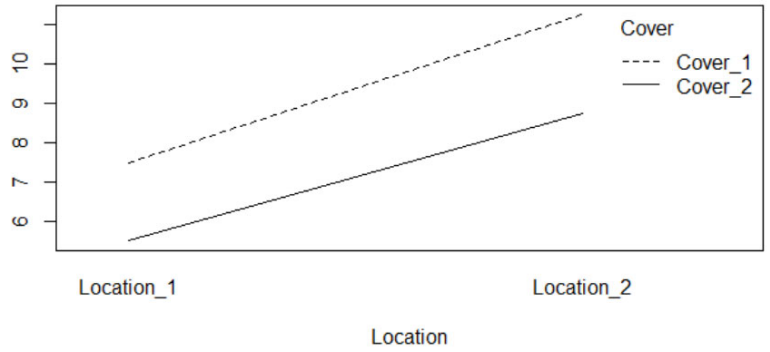
Effect of…	Conclusions from ANOVA	Conclusions from interaction plot
Cover	*Since p-value < 0.05 : "The magazine cover has a significant effect on sales"*	*"Sales are significantly higher when the cover is Cover 1, compared to when the Cover is Cover 2"*
Location	*Since p-value < 0.05 : "The location of the magazine has a significant effect on sales"*	*"Sales are significantly higher when the magazine is displayed in Location 2, compared to when the magazine is displayed in Location 1"*
Interaction		

```
> Two_way_anova_Magazines_2 = aov(Sales ~ Cover +
Location + Cover*Location, data=Magazines_2)
> summary(Two_way_anova_Magazines_2)
                Df Sum Sq Mean Sq F value  Pr(>F)
Cover            1  20.25   20.25   6.845 0.02254 *
Location         1  49.00   49.00  16.563 0.00155 **
Cover:Location   1   0.25    0.25   0.085 0.77624
Residuals       12  35.50    2.96
```



Effect of…	Conclusions from ANOVA	Conclusions from interaction plot
Cover	*Since p-value < 0.05 :* *"The magazine cover has a significant effect on sales"*	*"Sales are significantly higher when the cover is Cover 1, compared to when the Cover is Cover 2"*
Location	*Since p-value < 0.05 :* *"The location of the magazine has a significant effect on sales"*	*"Sales are significantly higher when the magazine is displayed in Location 2, compared to when the magazine is displayed in Location 1"*
Interaction	*Since p-value > 0.05 :* *"The interaction effect is not significant"*	-

```
> Two_way_anova_Magazines_2 = aov(Sales ~ Cover +
Location + Cover*Location, data=Magazines_2)
> summary(Two_way_anova_Magazines_2)
                Df Sum Sq Mean Sq F value  Pr(>F)
Cover            1  20.25   20.25   6.845 0.02254 *
Location         1  49.00   49.00  16.563 0.00155 **
Cover:Location   1   0.25    0.25   0.085 0.77624
Residuals       12  35.50    2.96
```



Effect of…	Conclusions from ANOVA	Conclusions from interaction plot
Cover	*Since p-value < 0.05 : "The magazine cover has a significant effect on sales"*	*"Sales are significantly higher when the cover is Cover 1, compared to when the Cover is Cover 2"*
Location	*Since p-value < 0.05 : "The location of the magazine has a significant effect on sales"*	*"Sales are significantly higher when the magazine is displayed in Location 2, compared to when the magazine is displayed in Location 1"*
Interaction	*Since p-value > 0.05 : "The interaction effect is not significant"*	-

My recommendation to the magazine editor:

If you want to sell more magazines:

Use Cover 1

and

Place it in Location 2 (high).

Let's practice with another dataset…

Dataset_3

Factor B: Location

Location 1

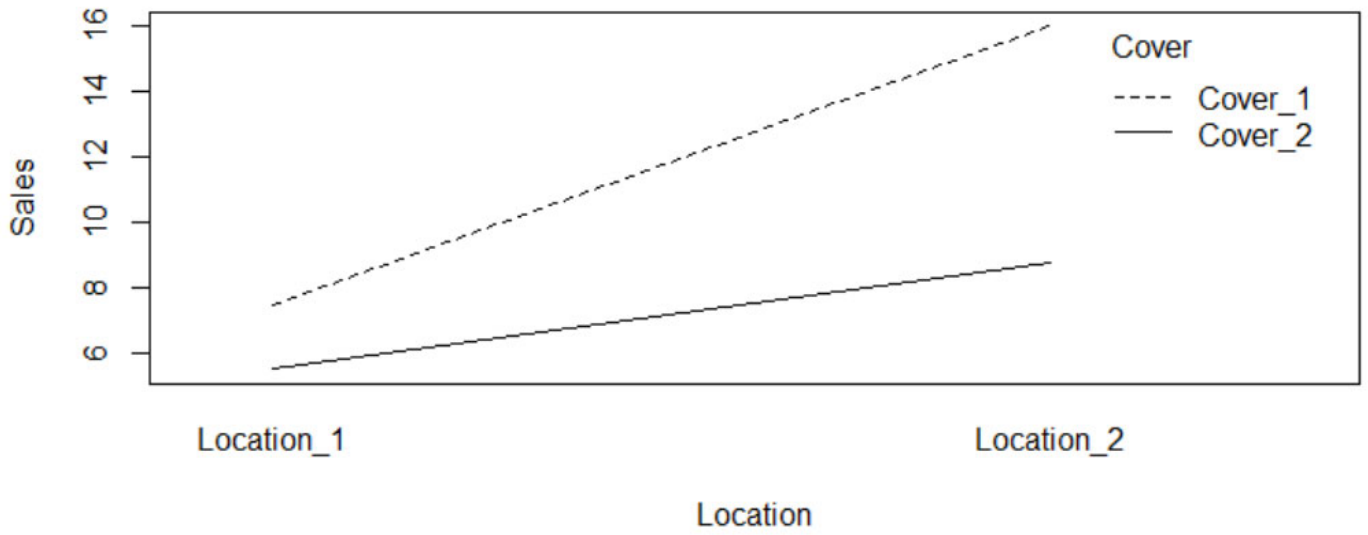
Location 2

Factor A: Cover	**Location 1**	**Location 2**
Cover 1 	6, 8	18, 15
	6, 10	12, 19
Cover 2 	5, 8	7, 10
	5, 4	10, 8

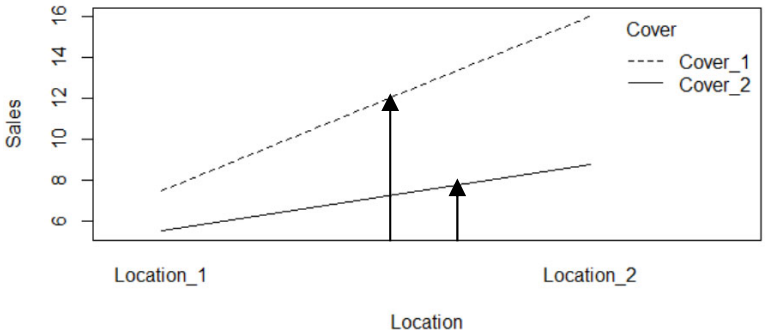
```
> Magazines_3 <-
read.table("C:/Users/vbv/Desktop/My_documents/Teaching/Teaching/
Statistical_Data_Analysis/2022_2023/Exercises_in_R/Magazines_Exa
mple_3.txt", header=TRUE)
> Magazines_3
     Cover   Location Sales
1  Cover_1 Location_1     6
2  Cover_1 Location_1     8
3  Cover_1 Location_1     6
4  Cover_1 Location_1    10
5  Cover_1 Location_2    18
6  Cover_1 Location_2    15
7  Cover_1 Location_2    12
8  Cover_1 Location_2    19
9  Cover_2 Location_1     5
10 Cover_2 Location_1     8
11 Cover_2 Location_1     5
12 Cover_2 Location_1     4
13 Cover_2 Location_2     7
14 Cover_2 Location_2    10
15 Cover_2 Location_2    10
16 Cover_2 Location_2     8
```

```
> Two_way_anova_Magazines_3 = aov(Sales ~ Cover + Location + Cover*Location,
data=Magazines_3)
> summary(Two_way_anova_Magazines_3)
               Df Sum Sq Mean Sq F value   Pr(>F)
Cover           1  85.56   85.56  18.093 0.001120 **
Location        1 138.06  138.06  29.194 0.000159 ***
Cover:Location  1  27.56   27.56   5.828 0.032664 *
Residuals      12  56.75    4.73
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
> interaction.plot(Magazines_3$Location, Magazines_3$Cover,
Magazines_3$Sales, xlab = "Location", ylab = "Sales", trace.label="Cover")
```

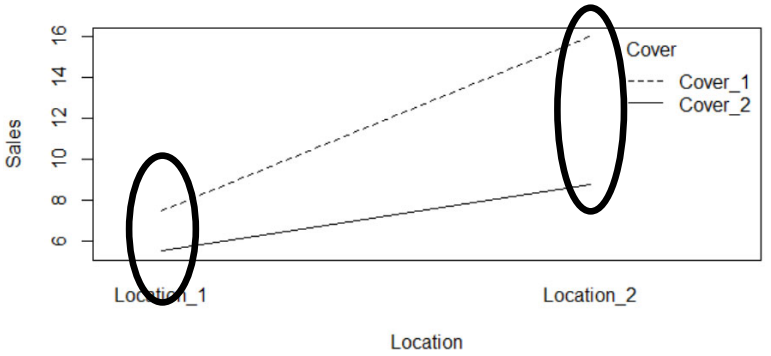


```
> Two_way_anova_Magazines_3 = aov(Sales ~ Cover +
Location + Cover*Location, data=Magazines_3)
> summary(Two_way_anova_Magazines_3)
                Df Sum Sq Mean Sq F value   Pr(>F)
Cover            1  85.56   85.56  18.093 0.001120 **
Location         1 138.06  138.06  29.194 0.000159 ***
Cover:Location   1  27.56   27.56   5.828 0.032664 *
Residuals       12  56.75    4.73
```



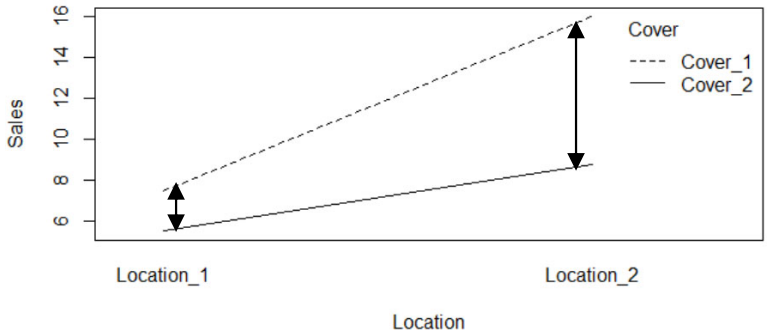
Effect of…	Conclusions from ANOVA	Conclusions from interaction plot
Cover	*Since p-value < 0.05 : "The magazine cover has a significant effect on sales"*	*"Sales are significantly higher when the cover is Cover 1, compared to when the Cover is Cover 2"*
Location		
Interaction		

```
> Two_way_anova_Magazines_3 = aov(Sales ~ Cover +
Location + Cover*Location, data=Magazines_3)
> summary(Two_way_anova_Magazines_3)
                Df Sum Sq Mean Sq F value   Pr(>F)
Cover            1  85.56   85.56  18.093 0.001120 **
Location         1 138.06  138.06  29.194 0.000159 ***
Cover:Location   1  27.56   27.56   5.828 0.032664 *
Residuals       12  56.75    4.73
```



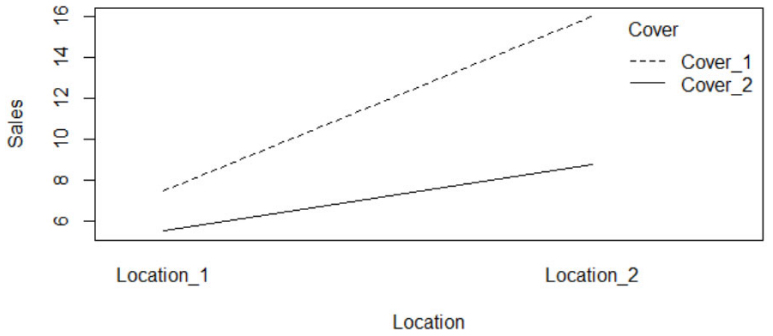
Effect of…	Conclusions from ANOVA	Conclusions from interaction plot
Cover	*Since p-value < 0.05 : "The magazine cover has a significant effect on sales"*	*"Sales are significantly higher when the cover is Cover 1, compared to when the Cover is Cover 2"*
Location	*Since p-value < 0.05 : "The location of the magazine has a significant effect on sales"*	*"Sales are significantly higher when the magazine is in Location 2, compared to when the magazine is in Location 1"*
Interaction		

```
> Two_way_anova_Magazines_3 = aov(Sales ~ Cover +
Location + Cover*Location, data=Magazines_3)
> summary(Two_way_anova_Magazines_3)
                Df Sum Sq Mean Sq F value   Pr(>F)
Cover            1  85.56   85.56  18.093 0.001120 **
Location         1 138.06  138.06  29.194 0.000159 ***
Cover:Location   1  27.56   27.56   5.828 0.032664 *
Residuals       12  56.75    4.73
```



Effect of…	Conclusions from ANOVA	Conclusions from interaction plot
Cover	*Since p-value < 0.05 :* *"*The magazine cover has a significant effect on sales"	*"Sales are significantly higher when the cover is Cover 1, compared to when the Cover is Cover 2"*
Location	*Since p-value < 0.05 :* *"The location of the magazine has a significant effect on sales"*	*"Sales are significantly higher when the magazine is in Location 2, compared to when the magazine is in Location 1"*
Interaction	*Since p-value < 0.05 :* *"The interaction effect is significant"* *In other words: "The effect of one factor depends on the level of the other factor"*	*"In general, Cover 1 results in more sales (compared to Cover 2), but most specially when magazine is displayed in Location 2"*

```
> Two_way_anova_Magazines_3 = aov(Sales ~ Cover +
Location + Cover*Location, data=Magazines_3)
> summary(Two_way_anova_Magazines_3)
                Df Sum Sq Mean Sq F value   Pr(>F)
Cover            1  85.56   85.56  18.093 0.001120 **
Location         1 138.06  138.06  29.194 0.000159 ***
Cover:Location   1  27.56   27.56   5.828 0.032664 *
Residuals       12  56.75    4.73
```



Effect of…	Conclusions from ANOVA	Conclusions from interaction plot
Cover	*Since p-value < 0.05 :* *"*The magazine cover has a significant effect on sales"	*"Sales are significantly higher when the cover is Cover 1, compared to when the Cover is Cover 2"*
Location	*Since p-value < 0.05 :* *"The location of the magazine has a significant effect on sales"*	*"Sales are significantly higher when the magazine is in Location 2, compared to when the magazine is in Location 1"*
Interaction	*Since p-value < 0.05 :* *"The interaction effect is significant"* *In other words: "The effect of one factor depends on the level of the other factor"*	*"In general, Cover 1 results in more sales (compared to Cover 2), but most specially when magazine is displayed in Location 2"*

My recommendation to the magazine editor:

If you want to sell more magazines:

Use Cover 1

and

Place the magazine in Location 2 (high).

The interaction effect makes this combination (Cover 1 with Location 2) proportionally MUCH better than any other.

Let's practice with another dataset…

Dataset_4

Factor B: Location

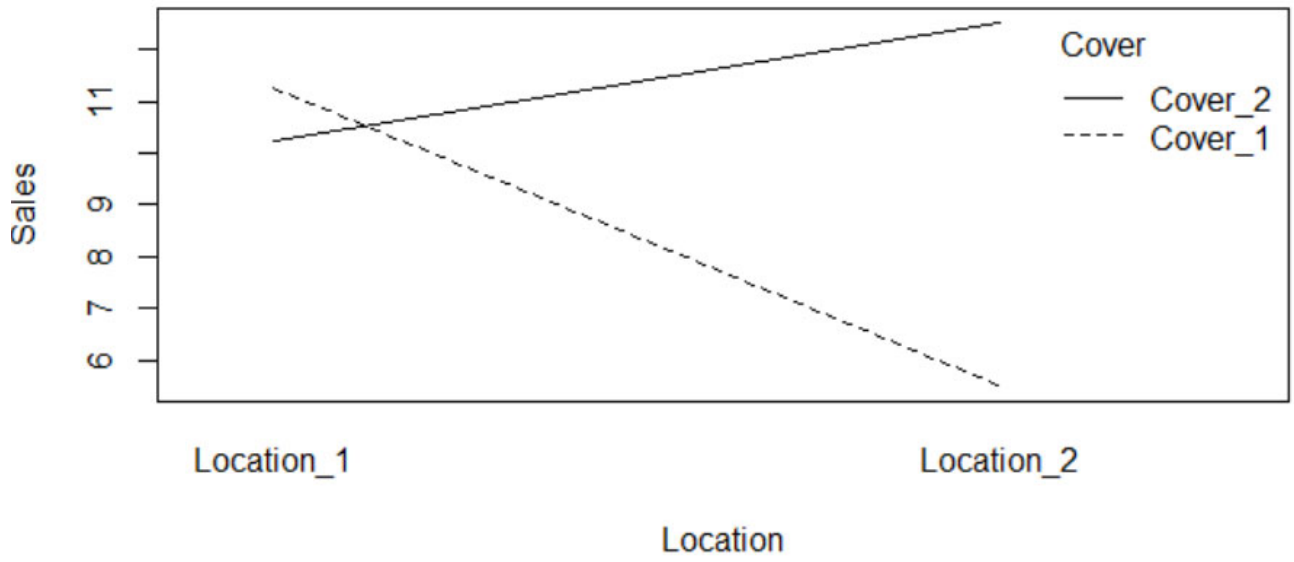
Location 1

Location 2

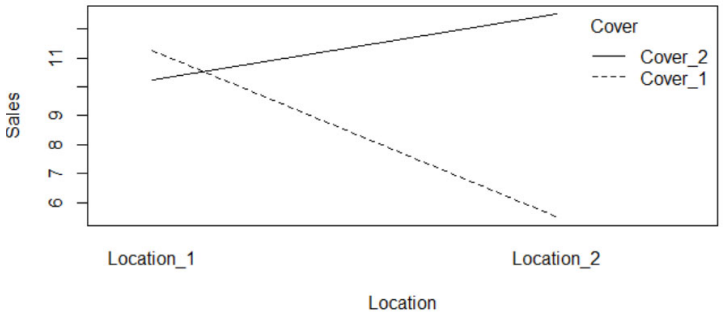
Factor A: Cover			
Cover 1 	10, 12	13, 10	5, 8	5, 4
Cover 2 	12, 9	11, 9	13, 14	11, 12

```
> Magazines_4 <-
read.table("C:/Users/vbv/Desktop/My_documents/Teaching/Teaching/
Statistical_Data_Analysis/2022_2023/Exercises_in_R/Magazines_Exa
mple_4.txt", header=TRUE)
> Magazines_4
     Cover   Location Sales
1  Cover_1 Location_1    10
2  Cover_1 Location_1    12
3  Cover_1 Location_1    13
4  Cover_1 Location_1    10
5  Cover_1 Location_2     5
6  Cover_1 Location_2     8
7  Cover_1 Location_2     5
8  Cover_1 Location_2     4
9  Cover_2 Location_1    12
10 Cover_2 Location_1     9
11 Cover_2 Location_1    11
12 Cover_2 Location_1     9
13 Cover_2 Location_2    13
14 Cover_2 Location_2    14
15 Cover_2 Location_2    11
16 Cover_2 Location_2    12
```

```
> Two_way_anova_Magazines_4 = aov(Sales ~ Cover + Location +
Cover*Location, data=Magazines_4)
> summary(Two_way_anova_Magazines_4)
               Df Sum Sq Mean Sq F value   Pr(>F)
Cover           1  36.00   36.00  15.709 0.001882 **
Location        1  12.25   12.25   5.345 0.039332 *
Cover:Location  1  64.00   64.00  27.927 0.000193 ***
Residuals      12  27.50    2.29
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> interaction.plot(Magazines_4$Location, Magazines_4$Cover,
Magazines_4$Sales, xlab = "Location", ylab = "Sales",
trace.label="Cover")
```

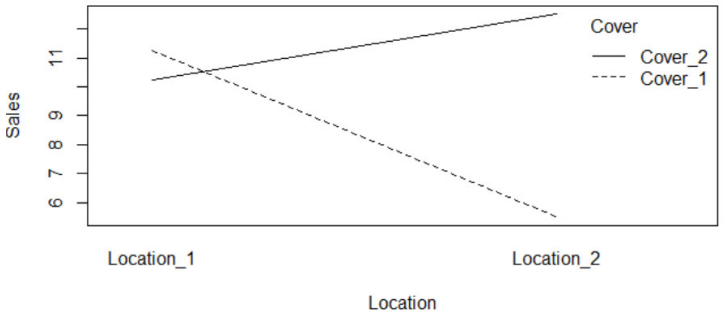


```
> Two_way_anova_Magazines_4 = aov(Sales ~ Cover +
Location + Cover*Location, data=Magazines_4)
> summary(Two_way_anova_Magazines_4)
                Df Sum Sq Mean Sq F value   Pr(>F)
Cover            1  36.00   36.00  15.709 0.001882 **
Location         1  12.25   12.25   5.345 0.039332 *
Cover:Location   1  64.00   64.00  27.927 0.000193 ***
Residuals       12  27.50    2.29
```



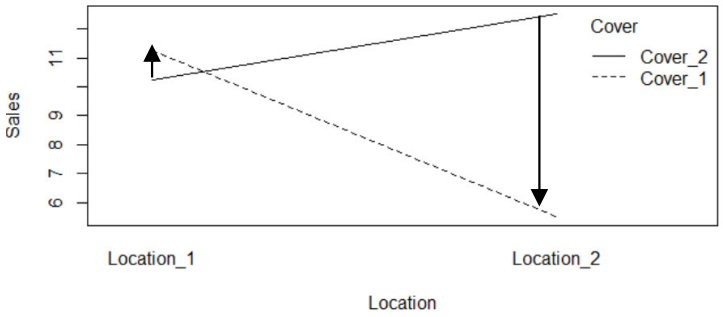
Effect of…	Conclusions from ANOVA	Conclusions from interaction plot
Cover	*Since p-value < 0.05 : "The magazine cover has a significant effect on sales"*	-
Location		
Interaction		

```
> Two_way_anova_Magazines_4 = aov(Sales ~ Cover +
Location + Cover*Location, data=Magazines_4)
> summary(Two_way_anova_Magazines_4)
                Df Sum Sq Mean Sq F value   Pr(>F)
Cover            1  36.00   36.00  15.709 0.001882 **
Location         1  12.25   12.25   5.345 0.039332 *
Cover:Location   1  64.00   64.00  27.927 0.000193 ***
Residuals       12  27.50    2.29
```



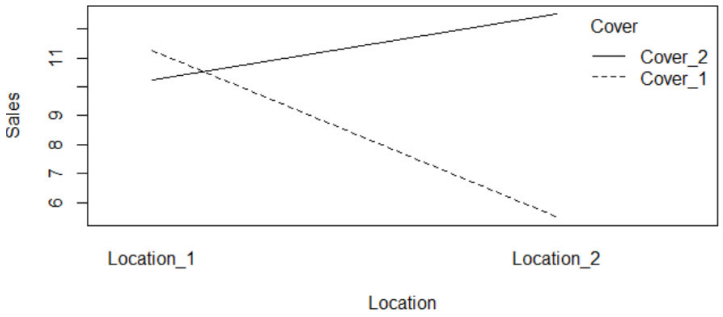
Effect of…	Conclusions from ANOVA	Conclusions from interaction plot
Cover	*Since p-value < 0.05 : "The magazine cover has a significant effect on sales"*	-
Location	*Since p-value < 0.05 : "The location of the magazine has a significant effect on sales"*	-
Interaction		

```
> Two_way_anova_Magazines_4 = aov(Sales ~ Cover +
Location + Cover*Location, data=Magazines_4)
> summary(Two_way_anova_Magazines_4)
                Df Sum Sq Mean Sq F value   Pr(>F)
Cover            1  36.00   36.00  15.709 0.001882 **
Location         1  12.25   12.25   5.345 0.039332 *
Cover:Location   1  64.00   64.00  27.927 0.000193 ***
Residuals       12  27.50    2.29
```



Effect of…	Conclusions from ANOVA	Conclusions from interaction plot
Cover	*Since p-value < 0.05 :* *"The magazine cover has a significant effect in sales"*	-
Location	*Since p-value < 0.05 :* *"The location of the magazine has a significant effect in sales"*	-
Interaction	*Since p-value < 0.05 :* *"The interaction effect is significant"* *"The effect of one factor depends on the level of the other factor"*	*"Sales are significantly affected by the combination of the type of cover and the location of the magazine"*

```
> Two_way_anova_Magazines_4 = aov(Sales ~ Cover +
Location + Cover*Location, data=Magazines_4)
> summary(Two_way_anova_Magazines_4)
                Df Sum Sq Mean Sq F value   Pr(>F)
Cover            1  36.00   36.00  15.709 0.001882 **
Location         1  12.25   12.25   5.345 0.039332 *
Cover:Location   1  64.00   64.00  27.927 0.000193 ***
Residuals       12  27.50    2.29
```



Effect of…	Conclusions from ANOVA	Conclusions from interaction plot
Cover	*Since p-value < 0.05 :* *"The magazine cover has a significant effect in sales"*	-
Location	*Since p-value < 0.05 :* *"The location of the magazine has a significant effect in sales"*	-
Interaction	*Since p-value < 0.05 :* *"The interaction effect is significant"* *"The effect of one factor depends on the level of the other factor"*	*"Sales are significantly affected by the combination of the type of cover and the location of the magazine"*

My recommendation to the magazine editor:

If you want to sell more magazines:

Use Cover 2 and place it in Location 2.

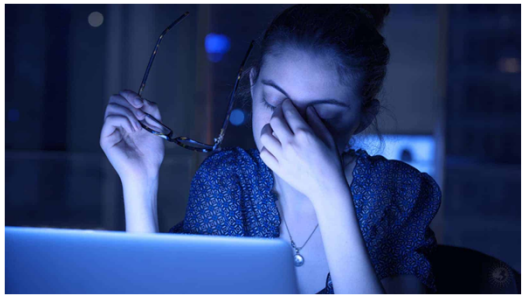
But, if the magazine has to be located in Location 1 (low), then do not use Cover 2. In that case, it is better to use Cover 1.

Analysis of variance: Two-way ANOVA

1. Recap: One-way Analysis of variance
2. What is “Two-Way Analysis of Variance”?
3. How can I interpret the results of a “Two-way ANOVA”?
    1. Example Magazines (Dataset 1)
    2. Example Magazines (Dataset 2)
    3. Example Magazines (Dataset 3)
    4. Example Magazines (Dataset 4)
4. In-class exercise

In-class exercise
Computer screen fatigue

Spending prolonged periods of time looking at a computer screen can cause **computer screen fatigue**.



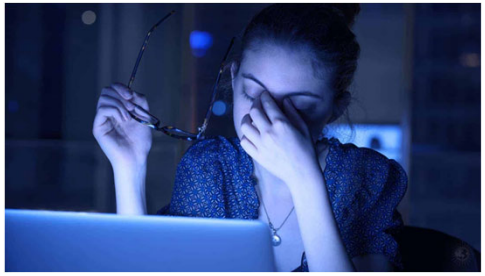
Different ways can help reducing computer screen fatigue, for example: Adjusting the brightness of the screen, increase blinking, use eyedrops to avoid dryness, etc.

We want to know whether “Taking breaks” and/or using “Blue-Light-Blocking Glasses” has a significant effect on reducing computer screen fatigue.

We ask 8 computer users to perform a task on the computer that lasted 4 hours, under the following conditions:

Factor A: Glasses	Factor B: Breaks — Yes	Factor B: Breaks — No
Blue-light blocking glasses 	3 2	3 5
Clear glasses 	4 3	7 6

Where numbers indicate computer fatigue (obtained from physiological measures) (the higher the number, the more fatigue)



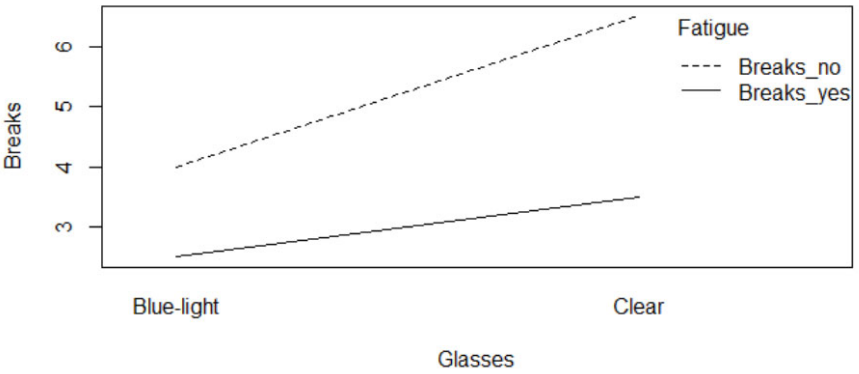
```
> Fatigue_data <-
read.table("C:/Users/vbv/Desktop/My_documents/Teaching/Teaching/Statistical_Data_Analysis/2022_2023/Exercises_in_R/Fatigue.txt",
header=TRUE)
> Fatigue_data
     Glasses     Breaks Fatigue
1 Blue-light Breaks_yes       3
2 Blue-light Breaks_yes       2
3 Blue-light  Breaks_no       3
4 Blue-light  Breaks_no       5
5      Clear Breaks_yes       4
6      Clear Breaks_yes       3
7      Clear  Breaks_no       7
8      Clear  Breaks_no       6

> Two_way_anova_Fatigue = aov(Fatigue ~ Glasses + Breaks +
Glasses*Breaks, data=Fatigue_data)
> summary(Two_way_anova_Fatigue)
               Df Sum Sq Mean Sq F value Pr(>F)
Glasses         1  6.125   6.125   7.000 0.0572 .
Breaks          1 10.125  10.125  11.571 0.0272 *
Glasses:Breaks  1  1.125   1.125   1.286 0.3202
Residuals       4  3.500   0.875
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
> Two_way_anova_Fatigue = aov(Fatigue ~ Glasses + Breaks +
Glasses*Breaks, data=Fatigue_data)
> summary(Two_way_anova_Fatigue)
                Df Sum Sq Mean Sq F value Pr(>F)
Glasses          1  6.125   6.125   7.000 0.0572 .
Breaks           1 10.125  10.125  11.571 0.0272 *
Glasses:Breaks   1  1.125   1.125   1.286 0.3202
Residuals        4  3.500   0.875
```

Considering a level of significance of 0.05, we have demonstrated that…

1. **The type of glasses has a significant effect on the fatigue.** — TRUE / FALSE
2. **Taking breaks has a significant effect on the fatigue.** — TRUE / FALSE
3. **There is a significant interaction effect.** — TRUE / FALSE

```
> Two_way_anova_Fatigue = aov(Fatigue ~ Glasses +
Breaks + Glasses*Breaks, data=Fatigue_data)
> summary(Two_way_anova_Fatigue)
               Df Sum Sq Mean Sq F value Pr(>F)
Glasses         1  6.125   6.125   7.000 0.0572 .
Breaks          1 10.125  10.125  11.571 0.0272 *
Glasses:Breaks  1  1.125   1.125   1.286 0.3202
Residuals       4  3.500   0.875
```



Considering a level of significance of 0.05, we have demonstrated that…

1. **In order to reduce fatigue, we should take breaks.** — TRUE / FALSE
2. **In order to reduce fatigue, we should wear blue-light blocking glasses.** — TRUE / FALSE
3. **If we are wearing clear glasses, it is particularly important to take breaks.** — TRUE / FALSE