

# Problem 1:

A research laboratory was developing a new compound for the relief of severe cases of hay fever. In an experiment with 36 volunteers, the amounts of the two active ingredients (A & B) in the compound were varied at three levels each. Randomization was used in assigning four volunteers to each of the nine treatments. The data on hours of relief can be found in the following .csv file: Fever.csv

[Assume all of the ANOVA assumptions are satisfied]

**1.1) State the Null and Alternate Hypothesis for conducting one-way ANOVA for both the variables 'A' and 'B' individually. [both statement and statistical form like  $H_0 = \mu$ ,  $H_a > \mu$ ]**

There are total 4 Volunteers, 9 treatments to each Volunteer.

Each treatment consists Different level of A and B ingredient.

Output or Dependent Variable is Relief in hours.

For Variable A,

$H_0 \Rightarrow \mu_1 = \mu_2 = \mu_3 = \mu_4$  (All volunteers shows same behaviour, Relief hours are independent to Level of A i.e Means are equal, Relief hours does not depends level of A)

$H_a \Rightarrow \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$  (All volunteers shows different behaviour i.e Atleast One mean is not equal, Relief hours depends on levels of A)

For Variable B,

$H_0 \Rightarrow \mu_1 = \mu_2 = \mu_3 = \mu_4$  (All volunteers shows same behaviour i.e Means are equal, Relief hours does not depends on levels of B)

$H_a \Rightarrow \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$  (All volunteers shows different behaviour with the differ in level of B i.e Atleast One mean is not equal, Relief hours depends on levels of B).

**1.2) Perform one-way ANOVA for variable 'A' with respect to the variable 'Relief'. State whether the Null Hypothesis is accepted or rejected based on the ANOVA results.**

One way ANOVA for variable 'A' with respect to the variable 'Relief'.

```
formula = 'Relief ~ C(A)'
model = ols(formula, df).fit()
aov_table = anova_lm(model)
print(aov_table)
```

	df	sum_sq	mean_sq	F	PR(>F)
C(A)	2.0	220.02	110.010000	23.465387	4.578242e-07
Residual	33.0	154.71	4.688182	NaN	NaN

For Variable A, Since the p value is less than the significance level (0.05), we can reject the null hypothesis and conclude that there is a difference in the mean of Relief for 4 volunteers. So we can conclude proportion of A variable matters to get relief.

**1.3) Perform one-way ANOVA for variable 'B' with respect to the variable 'Relief'. State whether the Null Hypothesis is accepted or rejected based on the ANOVA results.**

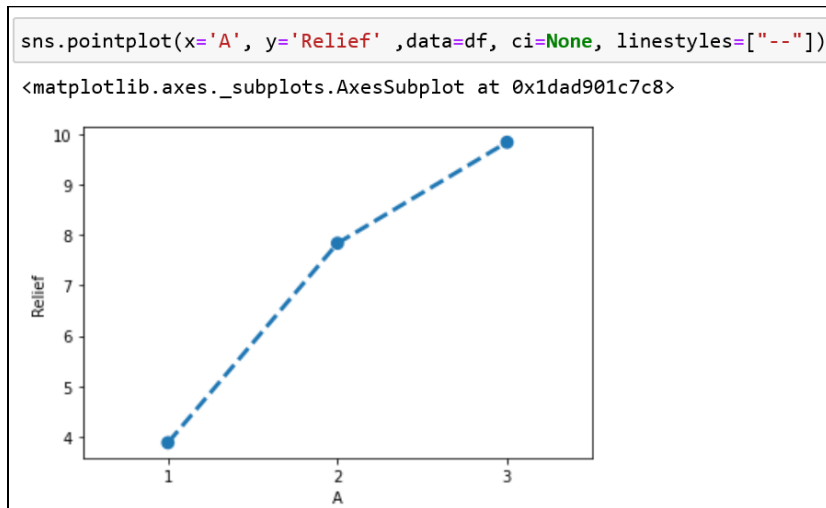
One way ANOVA for variable 'B' with respect to the variable 'Relief'.

```
formula = 'Relief ~ C(B)'  
model = ols(formula, df).fit()  
aov_table = anova_lm(model)  
print(aov_table)
```

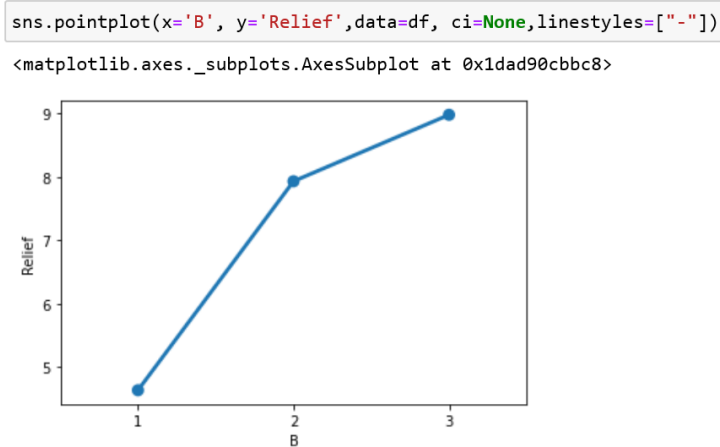
	df	sum_sq	mean_sq	F	PR(>F)
C(B)	2.0	123.66	61.830000	8.126777	0.00135
Residual	33.0	251.07	7.608182	NaN	NaN

For Variable B, Since the p value is less than the significance level (0.05), we can reject the null hypothesis and conclude that there is a difference in the mean of Relief for 4 volunteers. So we can conclude proportion of B variable matters to get relief.

**1.4) Analyse the effects of one variable on another with the help of an interaction plot. What is the interaction between the two treatments? [hint: use the 'pointplot' function from the 'seaborn' function]**



By looking at above point plot we can say that, Number of relief hours increases as we increase level of ingredient A.

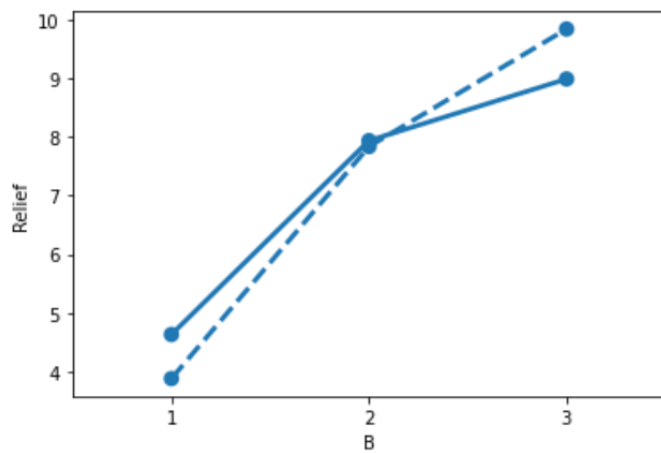


By looking at above point plot we can say that, Number of relief hours increases as we increase level of ingredient B.

Let's Compare the Variable A and B in one graph, Interaction Plot

```
sns.pointplot(x='A', y='Relief', data=df, ci=None, linestyle=["--"])
sns.pointplot(x='B', y='Relief', data=df, ci=None, linestyle=["--"])
```

<matplotlib.axes.\_subplots.AxesSubplot at 0x1dad9017c48>



In Above graph, We can see two lines, intersecting each other,

For Level 1, B shows greater effect than A--> Number of Relief hours is high for B ingredient.

For Level 2, Both same Almost Same effect on Volunteer --> Number of Relief hours is same for both ingredients.

For Level 3, A shows greater effect than B--> Number of Relief hours is high for A ingredient.

From the above results, it can be stated that in case of both (A & B) number of relief hours tend to increase with increase in the amount of ingredients.

**1.5) Perform a two-way ANOVA based on the different ingredients (variable 'A' & 'B' along with their interaction 'A\*B') with the variable 'Relief' and state your results.**

The primary purpose of a two-way ANOVA is to understand if there is an interaction between the two independent variables on the dependent variable.

Ho=>  $\mu_1=\mu_2=\mu_3=\mu_4$  (All volunteers shows same behaviour, Relief hours are independent to Level of A and B i.e Means are equal, Relief hours does not depends level of A and B)

Ha=>  $\mu_1\neq\mu_2\neq\mu_3\neq\mu_4$  (All volunteers shows different behaviour i.e Atleast One mean is not equal, Relief hours depends on levels of A and B Combination)

```
model1 = ols('Relief ~ C(A) + C(B) + C(A):C(B)', data=df).fit()
aov_table=anova_lm(model1,type=2)
print(aov_table)
```

	df	sum_sq	mean_sq	F	PR(>F)
C(A)	2.0	220.020	110.010000	1827.858462	1.514043e-29
C(B)	2.0	123.660	61.830000	1027.329231	3.348751e-26
C(A):C(B)	4.0	29.425	7.356250	122.226923	6.972083e-17
Residual	27.0	1.625	0.060185	NaN	NaN

By Performing two way Analysis, we can see the P values of A / B / A and B are less than 0.05, Null Hypothesis is Rejected.

Interaction of A and B affects No of Relief hours.

It suggests that both A and B affect Relief, Individually and Combining both.

***1.6) Mention the business implications of performing ANOVA for this particular case study.***

From the One way Anova test for Variable A, we got to Know That variable A is related to Relief hours. So While Producing Drug we have to consider A and Level of A. This Ingredient cannot be ignored.

From the One way Anova test for Variable B, we got to Know That variable A is related to Relief hours. So While Producing Drug we have to consider ingredient A and Level of A. This Ingredient cannot be ignored.

From the Two Way Anova we got to know that there is Interaction effect between ingredient A and ingredient B. These both might create wonders in case of Result as relief. So by Performing Further Tests we can Say to what proportion they must be added to get maximum results.

It is Mandatory to Look at all aspects of Ingerdient A and B, While designing the effective Drug. To take a Business Descision as which ingredient is making maximum effect or which is combination is better, one can perform Tukey honest Significant Difference.