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 file: [Fever.csv]

[Assume all of the `ANOVA` assumptions are satisfied]

1.1) State the `Null Hypothesis` and `Alternate Hypothesis` for conducting one-way `ANOVA` for both the variables `'A'` and `'B'` individually. [both statement and statistical form like `Ho=mu, Ha>mu`]

For conducting one-way `ANOVA` for variable `'A'`:

- The `Null Hypothesis` for `A` is that "the responses do not differ by the levels of factor `A`, while holding the levels of factor `B` constant and the interactions
- The alternate hypothesis is that the responses differ by the levels of factor `A`, while holding the levels of factor `B` constant and the interactions

In statistical form,

```
Ho : mu1 = mu2 = mu3
Ha : all population means are not equal
Where mu1, mu2, mu3 are the population means of factor A
```

For conducting one-way `ANOVA` for variable `'B'`:

- The `Null Hypothesis` for `B` is that the responses do not differ by the levels of factor `B`, while holding the levels of factor `A` constant and the interactions
- The alternate hypothesis is that the responses differ by the levels of factor `B`, while holding the levels of factor `A` constant and the interactions

In statistical form,

```
Ho :  mu1 = mu2 = mu3
Ha : all population means are not equal
Where mu1, mu2, mu3 are the population means of factor 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.

By performing one-way `ANOVA` for variable `A` with respect to the variable `Relief`, by using `model = ols(formula, data).fit()` , and by printing the `ANOVA` table, we got the output

	df	sum_sq	mean_sq	F	PR(>F)
Relief	1.0	13.604355	13.604355	44.494409	1.175871e-07
Residual	34.0	10.395645	0.305754	NaN	NaN

Since `p` value `< 0.05`, we reject the ``Null Hypothesis``. That means the responses differ by the levels of factor `A`, while holding the levels of factor `B` constant and the interactions.

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.

By performing one-way `ANOVA` for variable `'B'` with respect to the variable `'Relief'`, by using `model = ols(formula, data).fit()` , and by printing the `ANOVA` table, we got the output

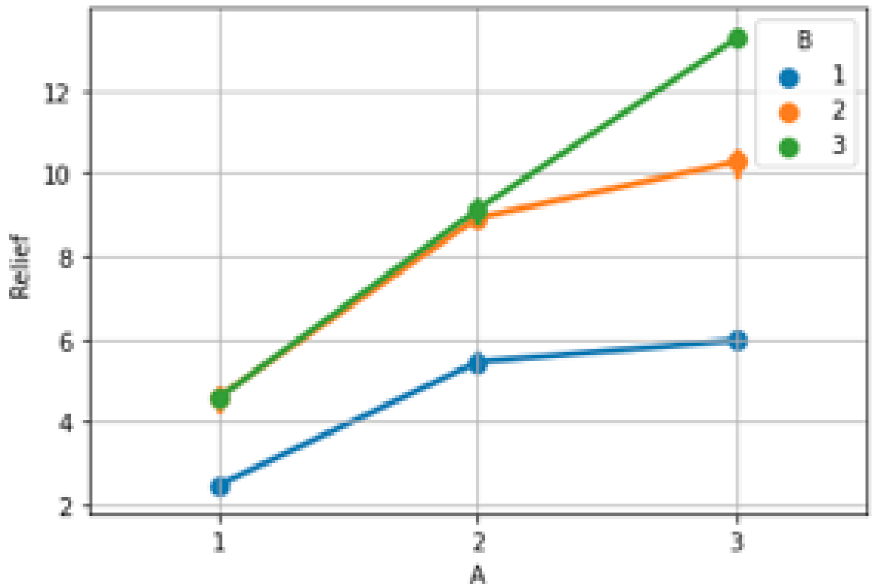
	df	sum_sq	mean_sq	F	PR(>F)
Relief	1.0	7.271475	7.271475	14.778958	0.000505
Residual	34.0	16.728525	0.492015	NaN	NaN

Since `p` value `< 0.05`, we reject the ``Null Hypothesis``. That means the responses differ by the levels of factor `B`, while holding the levels of factor `A` constant and the interactions.

1.4) Analyze 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 `point plot` function from the `seaborn` function]

I have used `point plot` in seaborn to plot the interaction between the two treatments



The plot suggests that there is interaction between the levels of ingredient A and ingredient B because the distance between the means across the three levels are not the same.

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.

By performing two-way ANOVA based on variable 'A' & 'B' along with their interaction with the variable 'Relief', by using `model = ols('Relief ~ A+B+(A*B)', data).fit()`, and by printing the ANOVA table, we got the output

	df	sum_sq	mean_sq	F	PR(>F)
A	1.0	212.415000	212.415000	308.976050	5.307332e-18
B	1.0	113.535000	113.535000	165.146510	3.529911e-14
A:B	1.0	26.780625	26.780625	38.954743	5.406597e-07
Residual	32.0	21.999375	0.687480	NaN	NaN

The Null Hypothesis for the interaction is that **there is no interaction between the levels of ingredient A and ingredient B**. The alternative hypothesis is **that there is interaction**. The test statistic is $F = 38.954743$ and the p value is less than 0.05. Therefore, at the alpha level of 0.05, we reject the Null Hypothesis and conclude that there is significant interaction between the levels of ingredient A and ingredient B.

It is possible for the interaction to be significant when the main effects are not significant. So, let's test the significance of the main effects. The Null Hypothesis for the main effect for A is that **the responses do not differ by the levels of factor `A`, while holding constant the levels of factor `B` and the interactions**. The `Null Hypothesis` for the main effect for `B` is that **the responses do not differ by the levels of factor `A`, while holding constant the levels of factor `A` and the interactions**. The test statistics for the main effects `A` and `B` are `F = 308.976050` and `F = 165.146510`, respectively. the `p` values are less than `0.05` for each. We reject the `Null Hypothesis` and conclude that the responses significantly differ across the levels of the two ingredients, while holding constant the other and the interactions.

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

`ANOVA` performs the test of equality of more than two population means by actually analyzing the variance. Here by considering that all the assumptions of `ANOVA` are satisfied, we performed one-way and two-way `ANOVA` to know the interaction between the levels of ingredient `A` and ingredient `B`. At a confidence interval of `95%`, we can conclude that there is significant interaction between the levels of ingredient `A` and ingredient `B` and also, we can conclude that the responses significantly differ across the levels of the two ingredients, while holding constant the other and the interactions.