FEVER- ANOVA ANALYSIS REPORT

Prepared by
Muralidhara N

ANOVA

1.1) State the Null and Alternate Hypothesis for conducting one-way ANOVA for both the variables 'A' and 'B' individually.

The one way ANOVA compares the means between the groups and determines whether any of the means are significantly different from each other.

The ANOVA states that at least two group means are significantly different from each other.

We are using one-way ANOVA to understand whether the relief hours differed based on active components of 'A' and 'B'

H_0: μ1 = μ2 = μ3 (The mean relief hours are equal with different levels of A)

H_a: Not all of the means are equal.

Significance Level α = 0.05

H_0: μ1 = μ2 = μ3 (The mean relief hours are equal with different levels of B)

H_a: Not all of the means are equal.

Significance Level α = 0.05

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.

```
formula = 'Relief ~ C(A)'
model = ols(formula, df_fever).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
```

Since the p value is less than the significance level, we can reject the null hypothesis and states that there is a difference in the mean of relief hours with different levels of A

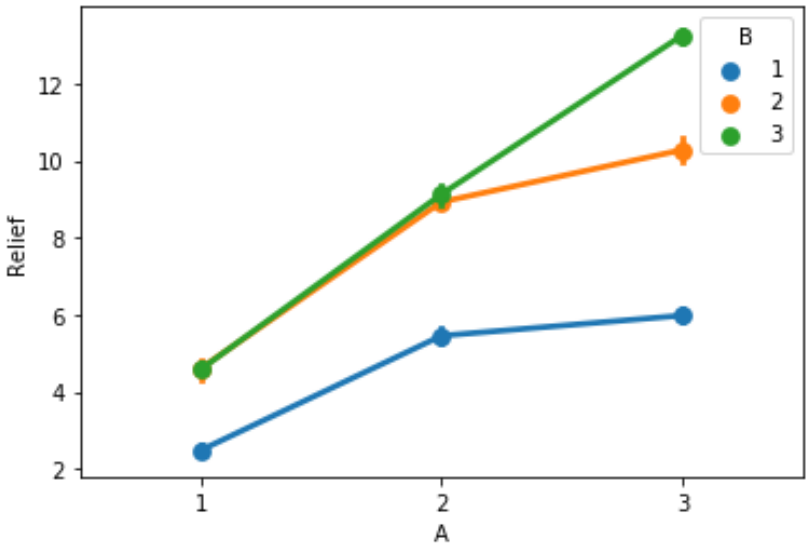
1.3) Perform one-way ANOVA for variable 'B' with respect to the variable 'Relief'. State whether the Null Hypothesis is accepted or rejected based on the ANOVA results.

```
formula = 'Relief ~ C(B)'
model = ols(formula, df_fever).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
```

Since the p value is less than the significance level, we can reject the null hypothesis and states that there is a difference in the mean of relief hours with different levels of B

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 'point plot' function from the 'sea born' function]



As we see from the interaction plots, there seems to be interaction amongst the two variables. We could understand there is overlapping between the plots and this clearly indicates interaction between variables A and B with respect to the treatments in reducing the relief hours of the volunteers who are affected by the hay fever.

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 result.

STEP 1:

NULL HYPOTHESIS

***H*0: The means of 'Relief' variable with respect to variable A and B is equal.**

ALTERNATE HYPOTHESIS

***H*1: At least one of the means of 'Relief' variable with respect variable A and B is unequal**

STEP 2

```
model=ols('Relief~C(A)+C(B)+C(A):C(B)',data=df_fever).fit()
aov_table=anova_lm(model)
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

STEP 3

Considering both the variable (A and B), A is a significant factor as the p value is <0.05 whereas B is also significant variable as p value of diet is >0.05

As A and B interaction is 6.972083e-17 which is < 0.05, there seems to be statistical interaction.

So as we know there is interaction between A and B this clearly indicates the compound has two active ingredients which is A and B which is also varied at three different levels, has significant effect in interaction with each other and interaction of A and B ingredients we can come to conclusion there is interaction which affects the relief hours of volunteers with the treatments.

STEP 4

There is significant interaction of relief hours with the treatments given ingredients A and B used at varied levels.

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

This business case problem is about developing a new compound for the relief of hay fever, the compound contains two active ingredients A and B that is varied at three different levels. A total of nine treatments and by using randomization four volunteers are assigned to each of nine treatments. The reason for performing ANOVA for this business case is that to find out the means are same or at least one pair of means are the same. (i.e.) would be the null and alternate hypothesis for this case. In this business case we have used to ONE-WAY ANOVA for A and B, We have also performed TWO-WAY ANOVA WITH INTERACTION EFFECT between the variables, the reason for using two-way is that we need a quantitative output value and you have two or more categorical variables in your case. The interaction plot has significantly made us understand that there is interaction between the variable A and B with respect to relief variable so performing two-way helps us to conclude that in a statistical way. Considering the interaction value which is less than P value which proves us that there is interaction between the variable A and B which significantly affects the relief hours of nine different treatments among the volunteers.