Problem Statement:

The staff of a service center for electrical appliances include three technicians who specialize in repairing three widely used electrical appliances by three different manufacturers. It was desired to study the effects of Technician and Manufacturer on the service time. Each technician was randomly assigned five repair jobs on each manufacturer's appliance and the time to complete each job (in minutes) was recorded. The data for this particular experiment is thus attached.

Exploratory Data Analysis:

	Technician	Manufacturer	Job	Service Time
0	1	1	1	62
1	1	1	2	48
2	1	1	3	63
3	1	1	4	57
4	1	1	5	69

Dataset has 3 variables Technician, Manufacturer & job, which has the different levels of Technician and Manufacturer and their respective Service Time.

Technician, Manufacturer & Job are categorical data types and Service Time is in float data type.

1. **State the Null and Alternate Hypothesis for conducting one-way ANOVA for both the variables ‘Manufacturer’ and ‘Technician’ individually.**

a) Variable:'Manufacturer' -

Null hypothesis states that the mean time to perform a work('Service Time') for all the Manufacturers are equal.

Alternative hypothesis states that there will be an effect of ‘Manufacturer’ on at least one of levels in Service Time. The mean time to perform a work('Service Time') for at least one category of Manufacturers are unequal.

H0 : µM1 = µM2 = µM3
H0 : The means between various manufacturers are equal

H1 : µM1 ≠ µM2 = µM3 or H1 : µM1 = µM2 ≠ µM3 or H1 : µM1 = µM3 ≠ µM2 or H1 : µM1 ≠ µM2 ≠ µM3
H1 : The means between various manufacturers are unequal (i.e)(Atleast one of the means b/w Manufacturers are unequal)

b) Variable:Technician -

Null hypothesis states that the mean time to perform a work('Service Time') for all the category of Technicians are equal.

Alternative hypothesis states that the mean time to perform a work('Service Time') for at least one category of Technicians are unequal.

H0 : **μT1 = μT2 = μT3**
H0 : The means between various Technician are equal

H1 : μT1 ≠ μT2 = μT3 or H1 : μT1 = μT2 ≠ μT3 or H1 : μT1 = μT3 ≠ μT2 or H1 : μT1 ≠ μT2 ≠ μT3
H1 : The means between various Technician are unequal (i.e)(Atleast one of means b/w Technicians are unequal)

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

First decide the level of significance:

The level of significance is defined as the probability of rejecting a null hypothesis by the test when it is really true, which is denoted as α. That is, P (Type I error) = α.
Confidence level: The level of significance 0.05 is related to the 95% confidence level.

Level of Significance α = 0.05

By performing one-way ANOVA, we got the below results:

	df	sum_sq	mean_sq	F	PR(>F)
Manufacturer	2.0	28.311111	14.155556	0.191029	0.826822
Residual	42.0	3112.266667	74.101587	NaN	NaN

Interpretation: The p-value obtained from ANOVA analysis for 'Manufacturer' is greater than α (0.05). We conclude that there is no difference in means between the different categories of Manufacturers with respect to Service Time. (i.e) All the means of Manufacturers with respect to the Service Time are equal.
If **p_value is less than α**, we have evidence to reject the null hypothesis since p value < Level of significance
If **p_value is greater than α**, we have fail to reject the null hypothesis since p value > Level of significance
Since **p_value > alpha_value**, here we *fail to reject the null hypothesis*.

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

Level of Significance α = 0.05

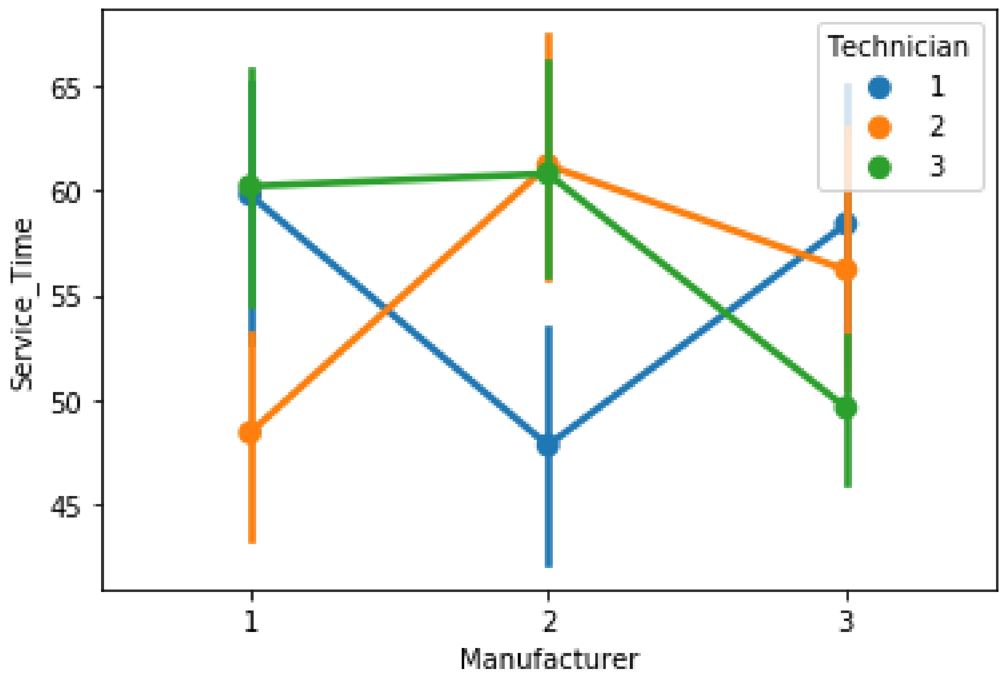
By performing one-way ANOVA, we got the below results:

	df	sum_sq	mean_sq	F	PR(>F)
Technician	2.0	24.577778	12.288889	0.16564	0.847902
Residual	42.0	3116.000000	74.190476	NaN	NaN

Interpretation: The P-value obtained from ANOVA analysis for 'Technician' is greater than α (0.05). We conclude that there is no difference in means between the different categories of Technicians with respect to Service Time. (i.e) All the means of Technicians with respect to the Service Time are equal.
If **p_value is less than α**, we have evidence to reject the null hypothesis since p value < Level of significance.
If **p_value is greater than α**, we fail reject the null hypothesis since p value > Level of significance.

Since **p_value > alpha_value**, here *we fail to reject the null hypothesis*.

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



Interaction effects occur when the effect of one variable depends on the value of another variable. Interaction effects are common in regression analysis, ANOVA, etc. Interaction effects indicate that a third variable influences the relationship between an independent and dependent variable. This type of effect makes the model more complex, but if the real world behaves this way, it is critical to incorporate it in your model. In this case, Service Time is the outcome which changes according to the various types of Technician and Manufacturer.
From the above interaction plot, since all the 3 lines are crossing each other, that means there is interaction effect of these 2 variables.

In this interaction plot, the lines are not parallel. This interaction effect indicates that the relationship between Technician and Manufacturer at some level

5. **Perform a two-way ANOVA based on the different ingredients (variable 'A' & 'B') and state your results.**

Level of Significance α = 0.05

By performing 2-way ANOVA, we got the below results:

	df	sum_sq	mean_sq	F	PR(>F)
Technician	2.0	24.577778	12.288889	0.236274	0.790779
Manufacturer	2.0	28.311111	14.155556	0.272164	0.763283
Technician:Manufacturer	4.0	1215.288889	303.822222	5.841487	0.000994
Residual	36.0	1872.400000	52.011111	NaN	NaN

Interpretation: The P-value obtained from ANOVA analysis for Manufacturer, Technician **are not statistically significant (P>0.05)** but the **interaction effect is significant p<0.05**. We conclude that type of levels of Manufacturer and Technician separately do not significantly affect the Service Time outcome but **Interaction of both** Manufacturer & Technician significantly **affects** the **Service Time** outcome. (i.e) For interaction term at least one of the means is different for the combination of Manufacturer and Technician w.r.t Service time.

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

The ANOVA test allows a comparison of more than two groups at the same time to determine whether a relationship exists between them. It allows for the analysis of multiple groups of data to determine the variability between samples and within samples.

In this case study, to create an effect of Technician and Manufacturer on the service time, ANOVA would help in the study of the varieties of 2 things namely Technician, Manufacturer, (Interaction effect) and its effects on Service_time. Variance analysis is important to assist with managing Service time by controlling levels or various types of Technician & Manufacturer and its corresponding interaction.