Aim:

The Aim of this study is to determine if the difference between conversion rate (as the number of users on which the intended effect was observed) in the control group and treatment group showed any statistical significance. Chi-square testing will be used to make the determination. Chi-square testing has various advantages that make it suitable in our case. Unlike t-test or z-test chi-square distribution does not assume normality and in our case where the outcome is binary, this test is the simplest to tell whether the difference observed (if any) is statistically significant. Confidence intervals will be determined for both the control and treatment group as well.

Introduction:

A/B testing is a common and reliable method for determining the effects of a new change on outcomes. It has a wide range of applications, including web page UX designs, medical trials, marketing campaigns, and more. Ultimately, A/B testing aims to determine whether a new protocol or procedure is actually responsible for causing the observed changes with statistical significance, or whether there is no meaningful difference between groups.

In our case, we use **synthetic, generic A/B testing data** consisting of two groups: **Control** and **Treatment**, with outcomes recorded as either **converted** or **not converted**.

- The **Control group** refers to participants who experienced the original, unchanged version of the product or setup.
- The **Treatment group** refers to participants who experienced a modified version, where something was intentionally changed.
- **Conversion** refers to participants performing the desired action — for example, signing up for a product, making a purchase, or clicking a button.

Dataset:

Description:

This dataset was synthetically generated using Python in response to an A/B testing simulation prompt. It consists of three columns: UserId, Group (Control or Treatment) and a Converted status column. Each row represents a unique user randomly assigned to either a control or treatment group. The Converted column indicates whether or not the user took the desired action (e.g., made a purchase or signed up). This dataset is not based on real user behavior but was created to simulate the structure and statistical challenges of real-world A/B testing scenarios.

Data Cleaning:

Data set was loaded into the jupyter notebook for analysis. Test was conducted to check for any missing and inconsistent values in the dataset. DataFrame had no missing values and the data type seems consistent.

Approach and Methodology:

The data frame was split into two subframes representing the Treatment and Control groups. Two functions — *conversion_counter* and *conversion_rate* — were written, which take a DataFrame as an argument and return the raw number of conversions and the percentage conversion, respectively. These functions were applied to the datasets to obtain both the conversion and non-conversion counts and rates for the Treatment and Control groups.

To compute the chi-square statistic, we first needed the expected conversion rates — that is, the overall conversion rate for the entire dataset regardless of group. This value is denoted as *t-converted* in the analysis notebook. Our null hypothesis is as follows: There are no statistically significant differences in the conversion rates between groups.

By applying the *t-converted* percentage to the size of the Control and Treatment groups, we obtained the expected number of conversions in each group. Subtracting this from the group size gives the expected number of non-conversions.

Finally the chi-square statistic is calculated using the following formula:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

χ^2 = the test statistic $\sum$ = the sum of

O = Observed frequencies E = Expected frequencies

The chi-square statistic calculated in our case ended up being : 1.167.

For a p-value of 5%, the chi-square statistic must be at least 3.84 to reject the null hypothesis. This threshold is determined by calculating the degrees of freedom and finding the corresponding critical value on the chi-square distribution table. Degrees of freedom represent the number of values that are free to vary independently in a dataset. In our case, each user can either be **converted** or **not converted**, and belong to either the **Treatment** or **Control** group.

The degrees of freedom for a chi-square test in a contingency table are calculated using the following formula:

$$\mathrm{df} = (\text{number of rows} - 1) \times (\text{number of columns} - 1)$$

$$\mathrm{df} = (2 - 1)(2 - 1) = 1$$

Number of rows and columns represent the group (treatment or control) and conversion outcomes (converted or non-converted).

Results:

From the chi square chart, for a p-value of 0.05, and degrees of freedom =1 the chi square statistic must be at least 3.841 to reject the null hypothesis. Since our chi-square statistic is 1.167 so we fail to reject the null hypothesis (that there is no statistically significant difference in the conversion of treatment or control group). The 95% confidence interval can be calculated as following for the treatment and the control group:

Control Group:

Conversion Rate

$$\hat{p}_C = \frac{120}{1016} = 0.1181$$

Standard Error

$$SE_C = \sqrt{\frac{0.1181 \cdot (1 - 0.1181)}{1016}} = \sqrt{\frac{0.1042}{1016}} = \sqrt{0.0001026} \approx 0.01013$$

z-score for 95% confidence

$$z = 1.96$$

Margin of Error

$$ME_C = 1.96 \cdot 0.01013 \approx 0.01985$$

95% Confidence Interval

$$CI_C = 0.1181 \pm 0.01985 = [0.0983, 0.13795]$$

Treatment group:

Conversion Rate

$$\hat{p}_T = \frac{122}{984} = 0.1240$$

Standard Error

$$SE_T = \sqrt{\frac{0.1240 \cdot (1 - 0.1240)}{984}} = \sqrt{\frac{0.1089}{984}} = \sqrt{0.0001106} \approx 0.01052$$

z-score for 95% confidence

$$z = 1.96$$

Margin of Error

$$ME_T = 1.96 \cdot 0.01052 \approx 0.02062$$

95% Confidence Interval

$$CI_T = 0.1240 \pm 0.02062 = [0.1034, 0.1446]$$

Control Range: 9.83% – 13.80%
Treatment Range: 10.34% – 14.46%

Assuming the experiment were repeated 100 times, 95 of those confidence intervals would be expected to contain the true conversion rate for the Control and Treatment groups.

Conclusion:

From this study, we conclude that the observed differences between the conversion rates of the Treatment and Control groups are **not statistically significant** and do not demonstrate a **causal relationship**.

Furthermore, the 95% confidence intervals for the Control and Treatment groups overlap, reinforcing the finding that the observed difference could likely be attributed to random variation rather than a real effect.