Hypothesis Testing Using Python:

Hypothesis testing is a statistical method used to make inferences or draw conclusions about a population based on a sample of data. By formulating and testing hypotheses, researchers can determine whether there is sufficient evidence to support or refute a given claim or assumption. In hypothesis testing, we typically start with a *null hypothesis* (H0), which represents a default or no-effect assumption, and an *alternative hypothesis* (H1 or Ha), which reflects the effect or difference we aim to detect.n.

Hypothesis testing Process:

Hypothesis testing is a core method in data science, allowing us to make data-driven inferences and decisions about populations based on sample data. Here's a structured process to follow for conducting hypothesis testing:

1. **Collect the Required Data**: Gather all necessary data to perform the hypothesis test.

2. **Define Hypotheses**: Establish the Null Hypothesis (H0) and Alternative Hypothesis (H1 or Ha).

3. **Set the Significance Level (α)**: Determine the probability threshold for rejecting the null hypothesis, which indicates the likelihood of a Type I error.

4. **Choose a Statistical Test**: Select an appropriate test, such as a t-test for comparing means, a chi-square test for categorical data, or ANOVA for comparing means across multiple groups.

5. **Execute the Test**: Apply the chosen test to your data.

6. **Interpret Results**: Analyze the p-value and draw conclusions based on the test results.

```
[1]:
import pandas as pd
from scipy.stats import ttest_ind
```

```
[3]:
df = pd.read_csv("C:\\Users\\hp\\Downloads\\website_ab_test.csv")

df.head()
```

[3]:

	Theme	Click Through Rate	Conversion Rate	Bounce Rate	Scroll_Depth	Age	Location	Session_Duration	Purchases	Added_to_Cart
0	Light Theme	0.054920	0.282367	0.405085	72.489458	25	Chennai	1535	No	Yes
1	Light Theme	0.113932	0.032973	0.732759	61.858568	19	Pune	303	No	Yes
2	Dark Theme	0.323352	0.178763	0.296543	45.737376	47	Chennai	563	Yes	Yes
3	Light Theme	0.485836	0.325225	0.245001	76.305298	58	Pune	385	Yes	No
4	Light Theme	0.034783	0.196766	0.765100	48.927407	25	New Delhi	1437	No	No

User Engagement Metrics Summary for Website Theme Evaluation:

The dataset examines the performance of two website themes, and our goal is to determine which theme performs better through hypothesis testing. Let's start with a summary of the dataset, including the number of records, the presence of missing values, and basic statistics for the numerical columns:

```
[6]:
# dataset summary
summary = {
    'Number of Records': df.shape[0],
    'Number of Columns': df.shape[1],
    'Missing Values': df.isnull().sum(),
    'Numerical Columns Summary': df.describe()
}
summary
```

```
[6]: {'Number of Records': 1000,
 'Number of Columns': 10,
 'Missing Values': Theme               0
 Click Through Rate    0
 Conversion Rate       0
 Bounce Rate           0
 Scroll_Depth          0
 Age                   0
 Location              0
 Session_Duration      0
 Purchases             0
 Added_to_Cart         0
 dtype: int64,
 'Numerical Columns Summary':        Click Through Rate  Conversion Rate  Bounce Rate  Scroll_Depth  \
 count         1000.000000      1000.000000  1000.000000   1000.000000
 mean             0.256048         0.253312     0.505758     50.319494
 std              0.139265         0.139092     0.172195     16.895269
 min              0.010767         0.010881     0.200720     20.011738
 25%              0.140794         0.131564     0.353609     35.655167
 50%              0.253715         0.252823     0.514049     51.130712
 75%              0.370674         0.373040     0.648557     64.666258
 max              0.499989         0.498916     0.799658     79.997108

               Age  Session_Duration
 count  1000.000000       1000.000000
 mean     41.528000        924.999000
 std      14.114334        508.231723
 min      18.000000         38.000000
 25%      29.000000        466.500000
 50%      42.000000        931.000000
 75%      54.000000       1375.250000
 max      65.000000       1797.000000  }
```

The dataset consists of 1,000 records and 10 columns, with no missing values. Here's a breakdown of the numerical columns:

- **Click Through Rate (CTR)**: This metric measures the percentage of clicks relative to views, ranging from 0.01 to 0.50, with an average CTR around 0.26. This suggests that, on average, 26% of visitors click on the content.

- **Conversion Rate**: This rate, similar in range to the CTR, spans from 0.01 to 0.50, with a mean of about 0.25. This implies that approximately 25% of visitors complete a desired action, like signing up or making a purchase.

- **Bounce Rate**: The bounce rate, which represents the percentage of visitors who leave without interacting further, varies from 0.20 to 0.80, with an average rate of 0.51. This suggests that, on average, 51% of visitors do not engage with additional content.

- **Scroll Depth**: This metric indicates how far down a page users scroll, ranging from 20.01% to nearly 80%, with an average depth of 50.32%. This shows a moderate level of engagement, with users scrolling about halfway through a page on average.

- **Age**: User age ranges between 18 and 65, with an average age of about 41.5 years, suggesting a diverse age range for website users.

- **Session Duration**: Session duration spans significantly, from as short as 38 seconds to around 1,800 seconds (or 30 minutes), with an average session lasting approximately 925 seconds (around 15 minutes). This indicates variable engagement levels, with an average session suggesting moderate interaction.

These summary statistics provide insight into user engagement and behavior metrics, which are key for evaluating website theme performance through hypothesis testing.

Comparative Analysis of Website Theme Performance Using Hypothesis Testing

Next, let's compare the performance of the two themes using key metrics such as the average Click Through Rate, Conversion Rate, Bounce Rate, and other relevant indicators. By analyzing these metrics for each theme, we can observe any notable differences in user engagement. Following this comparison, we'll conduct hypothesis testing to determine if any observed differences are statistically significant.

This approach allows us to:

1. **Assess User Engagement**: By comparing metrics like Click Through Rate and Conversion Rate, we can evaluate which theme attracts and converts users more effectively.

2. **Identify Bounce and Interaction Differences**: Looking at Bounce Rate and Scroll Depth helps us see if one theme keeps users engaged for longer.

3. **Validate Differences with Hypothesis Testing**: Hypothesis testing will help confirm whether the differences we observe are meaningful or if they could have occurred by chance, helping us make a data-driven decision about theme performance.

This process ensures a comprehensive and statistically backed evaluation of each theme's effectiveness.

```
[11]:
# Convert relevant columns to numeric
df['Conversion Rate'] = pd.to_numeric(df['Conversion Rate'], errors='coerce')

# Handle NaN values if necessary
df.dropna(subset=['Conversion Rate'], inplace=True)

# Group by 'Theme' and calculate mean values for the metrics
theme_performance = df.groupby('Theme').mean(numeric_only=True)

# Sort the data by conversion rate for a better comparison
theme_performance_sorted = theme_performance.sort_values(by='Conversion Rate', ascending=False)

# Display the sorted DataFrame
theme_performance_sorted
```

[11]:

Theme	Click Through Rate	Conversion Rate	Bounce Rate	Scroll_Depth	Age	Session_Duration
Light Theme	0.247109	0.255459	0.499035	50.735232	41.734568	930.833333
Dark Theme	0.264501	0.251282	0.512115	49.926404	41.332685	919.482490

The comparison of average performance metrics between the Light Theme and Dark Theme provides the following insights:

- **Click Through Rate (CTR):** The Dark Theme shows a slightly higher average CTR of 0.2645, while the Light Theme has a CTR of 0.2471.

- **Conversion Rate:** The Light Theme outperforms the Dark Theme with a marginally higher average Conversion Rate of 0.2555 compared to 0.2513.

- **Bounce Rate:** Users exhibit a higher Bounce Rate with the Dark Theme (0.5121) than with the Light Theme (0.4990).

- **Scroll Depth:** On average, users scroll further in the Light Theme (50.74%) compared to the Dark Theme (49.93%).

- **Age:** The average age of users is comparable across both themes, with the Light Theme averaging about 41.73 years and the Dark Theme at approximately 41.33 years.

- **Session Duration:** Users engaged with the Light Theme spend slightly more time on the site, averaging 930.83 seconds, compared to 919.48 seconds for the Dark Theme.

The Dark Theme seems to attract clicks more effectively, but the Light Theme outperforms in terms of conversion rates, bounce rates, and user engagement metrics like scroll depth and session duration

Getting Started with Hypothesis Testing

We will set a significance level (alpha) of 0.05 for our hypothesis testing. This means we will consider a result statistically significant if the p-value from our test is below 0.05.

Next, we will conduct hypothesis testing to compare the Conversion Rates between the Light Theme and Dark Theme. Our hypotheses are defined as follows:

- **Null Hypothesis (H0):** There is no difference in Conversion Rates between the Light Theme and Dark Theme.

- **Alternative Hypothesis (Ha):** There is a difference in Conversion Rates between the Light Theme and Dark Theme.

To compare the means of these two independent samples, we will use a two-sample t-test. Let's proceed with the test.

Hypothesis testing based on the Conversion Rate between the Light Theme and Dark Theme

[12]:
```
# extracting conversion rates for both themes
conversion_rates_light = df[df['Theme'] == 'Light Theme']['Conversion Rate']
conversion_rates_dark = df[df['Theme'] == 'Dark Theme']['Conversion Rate']

# performing a two-sample t-test
t_stat, p_value = ttest_ind(conversion_rates_light, conversion_rates_dark, equal_var=False)

t_stat, p_value
```

[12]:
```
(0.4748494462782632, 0.6349982678451778)
```

This code performs the following functions:

1. **Extracts Conversion Rates**: It retrieves the conversion rates for users who interacted with the Light Theme and the Dark Theme from a DataFrame (df).

2. **Conducts a Two-Sample T-Test**: It performs a statistical comparison (two-sample t-test) to evaluate if there is a significant difference between the average conversion rates of the two themes. The equal_var=False parameter indicates that the two samples may have different variances.

3. **Returns Statistics**: It outputs the t-statistic (t_stat) and p-value (p_value), which are used to assess the significance of the difference between the two groups' conversion rates.

The result of the two-sample t-test gives a p-value of approximately 0.635. Since this p-value is much greater than our significance level of 0.05, we do not have enough evidence to reject the

null hypothesis. Therefore, we conclude that there is no statistically significant difference in Conversion Rates between the Light Theme and Dark Theme based on the data provided.

Hypothesis testing based on the Click Through Rate (CTR)

```
[13]:
# extracting click through rates for both themes
ctr_light = df[df['Theme'] == 'Light Theme']['Click Through Rate']
ctr_dark = df[df['Theme'] == 'Dark Theme']['Click Through Rate']

# performing a two-sample t-test
t_stat_ctr, p_value_ctr = ttest_ind(ctr_light, ctr_dark, equal_var=False)

t_stat_ctr, p_value_ctr
```

```
[13]:
(-1.9781708664172253, 0.048184353371010704)
```

The two-sample t-test for Click Through Rate (CTR) between the Light Theme and Dark Theme results in a p-value of about 0.048. Since this p-value is slightly below our significance level of 0.05, it suggests a statistically significant difference in CTRs between the two themes. This indicates that the Dark Theme likely has a higher CTR based on the direction of the test statistic.

Hypothesis Testing based on two other metrics: bounce rate and scroll depth

```
# extracting bounce rates for both themes
bounce_rates_light = df[df['Theme'] == 'Light Theme']['Bounce Rate']
bounce_rates_dark = df[df['Theme'] == 'Dark Theme']['Bounce Rate']

# performing a two-sample t-test for bounce rate
t_stat_bounce, p_value_bounce = ttest_ind(bounce_rates_light, bounce_rates_dark, equal_var=False)

# extracting scroll depths for both themes
scroll_depth_light = df[df['Theme'] == 'Light Theme']['Scroll_Depth']
scroll_depth_dark = df[df['Theme'] == 'Dark Theme']['Scroll_Depth']

# performing a two-sample t-test for scroll depth
t_stat_scroll, p_value_scroll = ttest_ind(scroll_depth_light, scroll_depth_dark, equal_var=False)
t_stat_scroll,p_value_scroll
```

```
[16]:
(0.7562277864140986, 0.4496919249484911)
```

```
[17]:
# creating a table for comparison
comparison_table = pd.DataFrame({
    'Metric': ['Click Through Rate', 'Conversion Rate', 'Bounce Rate', 'Scroll Depth'],
    'T-Statistic': [t_stat_ctr, t_stat, t_stat_bounce, t_stat_scroll],
    'P-Value': [p_value_ctr, p_value, p_value_bounce, p_value_scroll]
})

comparison_table
```

[17]:

	Metric	T-Statistic	P-Value
0	Click Through Rate	-1.978171	0.048184
1	Conversion Rate	0.474849	0.634998
2	Bounce Rate	-1.201888	0.229692
3	Scroll Depth	0.756228	0.449692

Click Through Rate: The test reveals a statistically significant difference, with the Dark Theme likely performing better (P-Value = 0.048).

Conversion Rate: No statistically significant difference was found (P-Value = 0.635).

Bounce Rate: There's no statistically significant difference in Bounce Rates between the themes (P-Value = 0.230).

Scroll Depth: Similarly, no statistically significant difference is observed in Scroll Depths (P-Value = 0.450).

In summary, while the two themes perform similarly across most metrics, the Dark Theme has a slight edge in terms of engaging users to click through. For other key performance indicators like Conversion Rate, Bounce Rate, and Scroll Depth, the choice between a Light Theme and a Dark Theme does not significantly affect user behaviour according to the data provided.