



A/B Testing for User Engagement: Light Theme vs. Dark Theme

Overview

An online bookstore aims to enhance its website design to boost user engagement and drive more book purchases. The platform currently offers users two theme options: "Light Theme" (existing design) and "Dark Theme" (new design).

To determine which theme leads to higher conversion rates, the bookstore's data science team designed an A/B testing experiment. This project involves analyzing user interaction data to evaluate the performance of both themes and guide the business in making data-informed design decisions.

Dataset: [Kaggle - Light Theme and Dark Theme: Case Study](#)

Key columns:

- **Theme:** Indicates which version of the website (Light or Dark) the user interacted with.

- **Click Through Rate (CTR):** Proportion of users who clicked on any links or buttons during their session — a measure of engagement.
- **Conversion Rate:** Percentage of users who signed up for an account after their first visit.
- **Bounce Rate:** Percentage of users who exited the website after viewing only one page — a key indicator of disinterest or poor user experience.
- **Scroll Depth:** How far down users scrolled on the website page, used to gauge content consumption and interaction depth.
- **Age:** The user's age, useful for demographic analysis.
- **Location:** The user's geographical location.
- **Session Duration:** The total time a user spent on the website during a session.
- **Added_to_Cart:** Indicates whether the user added any books to their shopping cart (Yes/No).
- **Purchases:** Whether the user completed a purchase during their session (Yes/No).

Tools & Techniques

- **Tools:** Python, Google Colab, Jupyter Notebook
- **Libraries:** Pandas (Data manipulation and cleaning), Matplotlib & Seaborn (Data visualization), SciPy (Hypothesis testing)

Methodology

Understanding the Data:

The dataset captures user interactions on an **online bookstore** website that offers two design themes: **Light** and **Dark**. Each row in the dataset represents a unique user session and includes several behavioral and demographic metrics relevant to analyzing user engagement and purchase behavior:

```

# Import libraries
import pandas as pd
import plotly.express as px
import plotly.graph_objects as go
import plotly.io as pio
from statsmodels.stats.proportion import proportions_ztest
from scipy import stats
pio.templates.default = "plotly_white"

# Upload file
file_path = '/content/drive/MyDrive/Data Analysis Projects/website_ab_test.csv'
data = pd.read_csv(file_path)
data.head()

data.info()

```

Exploratory Data Analysis (EDA)

1. Data Cleaning & Wrangling:

Before diving into analysis, the dataset was carefully cleaned to ensure accuracy and consistency:

- **Missing Values:** No missing values detected across any columns.
- **Data Types:** All columns have appropriate data types.

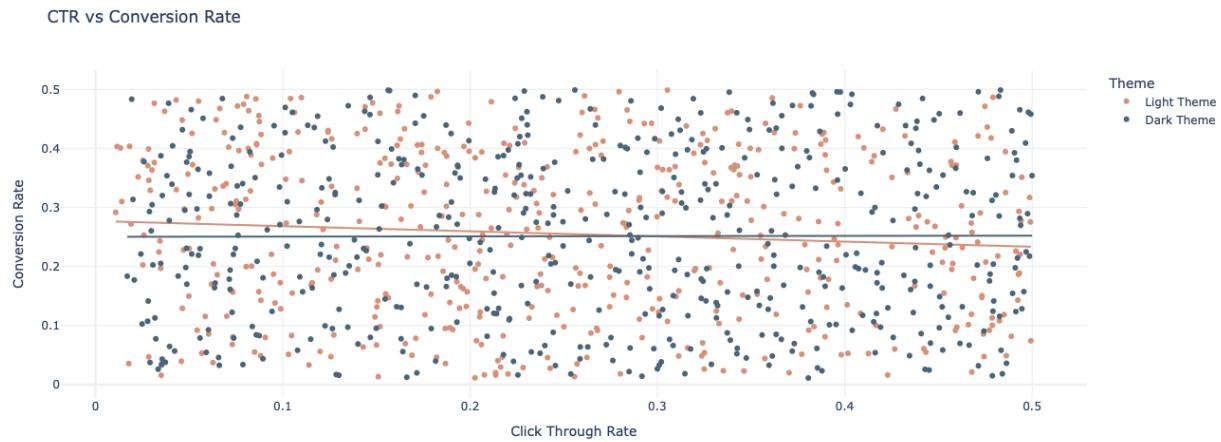
2. Correlation between Click-Through Rate and Conversion Rate

```

# Scatter plot for Click Through Rate and Conversion Rate
theme_colors = {'Light Theme': '#D8927A', 'Dark Theme': '#4D6879'}
fig = px.scatter(data, x='Click Through Rate',
                  y='Conversion Rate', color='Theme',
                  color_discrete_map=theme_colors,

```

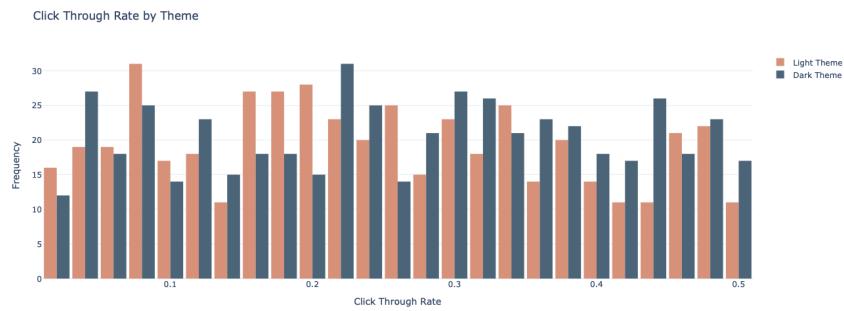
```
title='CTR vs Conversion Rate', trendline='ols')
fig.show()
```



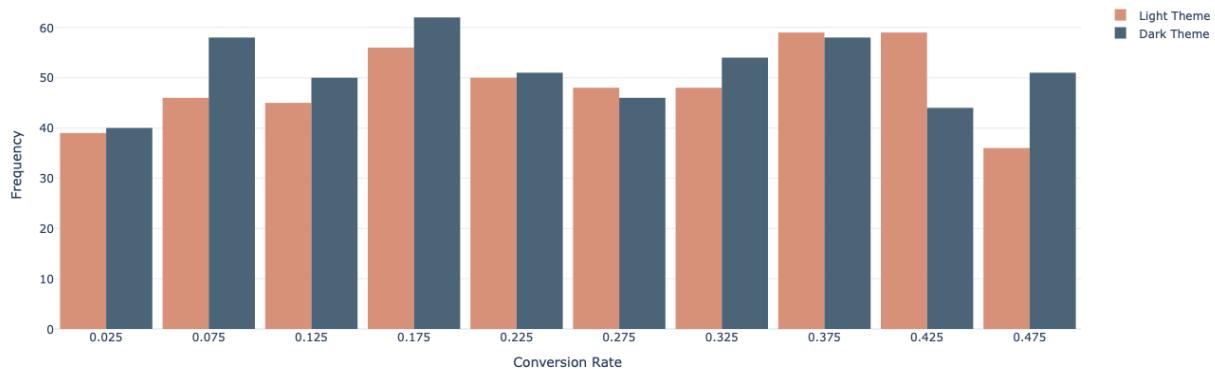
Scatter plots revealed a **weak correlation** for both themes.

3. Distribution Insights

CTR and Conversion Rate histograms showed slight performance differences. Although there's not much difference, the conversion rate of the dark theme is slightly better than the light theme.

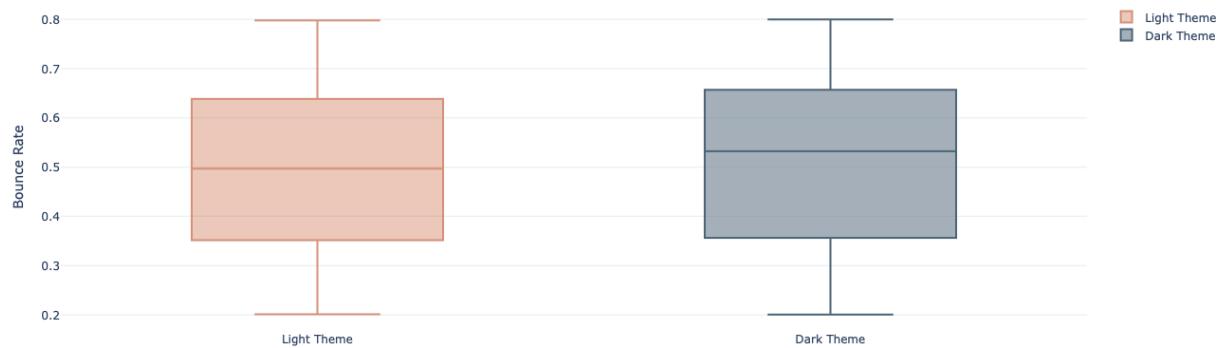


Conversion Rate by Theme

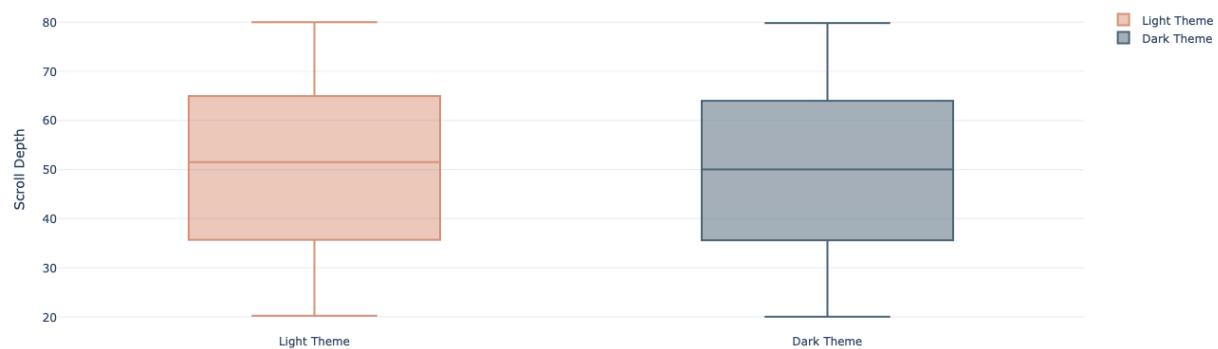


Bounce Rate and Scroll Depth were slightly better for the light theme.

Bounce Rate by Theme



Scroll Depth by Theme



A/B Testing for Purchases (Conversion Rate)

To evaluate whether the **Light Theme** or **Dark Theme** led to more purchases, I performed a **two-sample proportion z-test**, which is commonly used to compare conversion rates between two groups.

1. **Calculated conversions** (number of users who purchased) and total users for each theme.
2. **Computed conversion rates:**
 - Light Theme: $\text{conversion} = \text{purchases} / \text{total users}$
 - Dark Theme: $\text{conversion} = \text{purchases} / \text{total users}$
3. Used `proportions_ztest()` to determine if the difference in conversion rates is statistically significant.
 - Conducted a two-proportion z-test
 - Set significance level at 0.05
 - Null Hypothesis (H_0): There is **no difference** in conversion rates between the Light Theme and Dark Theme.
 - Alternative Hypothesis (H_1): There **is a difference** in conversion rates between the two themes.

```
# A/B testing for Purchases
light_theme_conversions = light_theme_data[light_theme_data['Purchases'] == 'Y']
light_theme_total = light_theme_data.shape[0]

dark_theme_conversions = dark_theme_data[dark_theme_data['Purchases'] == "N"]
dark_theme_total = dark_theme_data.shape[0]

conversion_counts = [light_theme_conversions, dark_theme_conversions]
sample_sizes = [light_theme_total, dark_theme_total]

light_theme_conversion_rate = light_theme_conversions / light_theme_total
dark_theme_conversion_rate = dark_theme_conversions / dark_theme_total
```

```
# Perform two-sample proportion test
zstat, pval = proportions_ztest(conversion_counts, sample_sizes)
print("Light Theme Conversion Rate:", light_theme_conversion_rate)
print("Dark Theme Conversion Rate:", dark_theme_conversion_rate)
print("A/B Testing - z-statistic:", zstat, " p-value:", pval)
```

Results:

- Light Theme Conversion Rate: 0.5308641975308642
- Dark Theme Conversion Rate: 0.5038910505836576
- A/B Testing: z-statistic: 0.8531246206222649; p-value: 0.39359019934127804

The p-value is well above the 0.05 significance level, meaning there is no strong evidence to reject the null hypothesis, thus we cannot confidently say that one theme performs better than the other in driving conversions.

A/B Testing for Session Duration

To determine if users engage longer with one theme over the other, I conducted a two-sample t-test comparing the average session durations for the Light and Dark themes.

- Calculated the average session duration for each group.
- Ran a two-sample t-test
- Null Hypothesis (H_0): The average session duration is the same for both themes.
- Alternative Hypothesis (H_1): The average session duration is different between the themes.

```
light_theme_session_duration = light_theme_data['Session_Duration']
dark_theme_session_duration = dark_theme_data['Session_Duration']
```

```
# Calculate the average session duration for both themes
```

```

light_theme_avg_duration = light_theme_session_duration.mean()
dark_theme_avg_duration = dark_theme_session_duration.mean()

# Print the average session duration for both themes
print("Light Theme Average Session Duration:", light_theme_avg_duration)
print("Dark Theme Average Session Duration:", dark_theme_avg_duration)

# Perform two-sample t-test for session duration
tstat, pval = stats.ttest_ind(light_theme_session_duration, dark_theme_session_du
print("A/B Testing for Session Duration - t-statistic:", tstat, " p-value:", pval)

```

Results:

- Light Theme Average Session Duration: 930.8333333333334
- Dark Theme Average Session Duration: 919.4824902723735
- A/B Testing for Session Duration: t-statistic: 0.3528382474155483; p-value: 0.7242842138292167

The difference in session duration is **not statistically significant**. The small variation observed is likely due to **random chance**, so we cannot conclude that either theme encourages longer user engagement.

Conclusion

Based on the A/B test results, there is no statistically significant difference between the two themes in terms of key performance metrics such as conversion rate, session duration, and click-through rate. While the dark theme shows a slightly higher conversion rate, the light theme performs marginally better in bounce rate and scroll depth. However, since these differences are minor and likely due to random variation rather than a true impact of the theme, we cannot confidently conclude that one theme is significantly better than the other.

Code & Visualizations: [Google Colab](#)