Hypothesis Testing Simulation

Objective: Simulate A/B testing to understand the effects of different variables on user engagement.

User response data for a web page is simulated, for the purpose of the simulation the assumption is of an e-commerce website testing a design change for instance changing the size of the buy now button For this there a two versions of the button one utilized in the control dataset and another in the treatment

The data is randomly generated using numpy in a function, with a seed value utilized to keep the generated data consistent between runs, a random probability is also generated an utilized with the function, with the time spent in the site kept within the range of 5 - 420min a figure gotten from a report done on average time spent on e-commerce website

The resulting data was clean for irregularities resulting from the mode of generation

The project objectives
we aimed to determine whether the new design led to a change in:

1. **click-through rate (CTR)**.
2. **conversion rate**.
3. **time spent** on the site in minutes.

Dataset Overview

The dataset consisted of the following key columns:

- ***user_id***: Unique identifier for each user.
- ***id***: Identifies whether the user was part of the **control** or **treatment** group.
- ***click***: Boolean value indicating whether a user clicked on the website.
- ***conversion***: Boolean value indicating whether a user converted (e.g., made a purchase or completed a goal).
- ***time_spent (min):*** Time (in minutes) that a user spent on the website.

The dataset included **[2000] rows** in total, evenly distributed across the control and treatment groups.

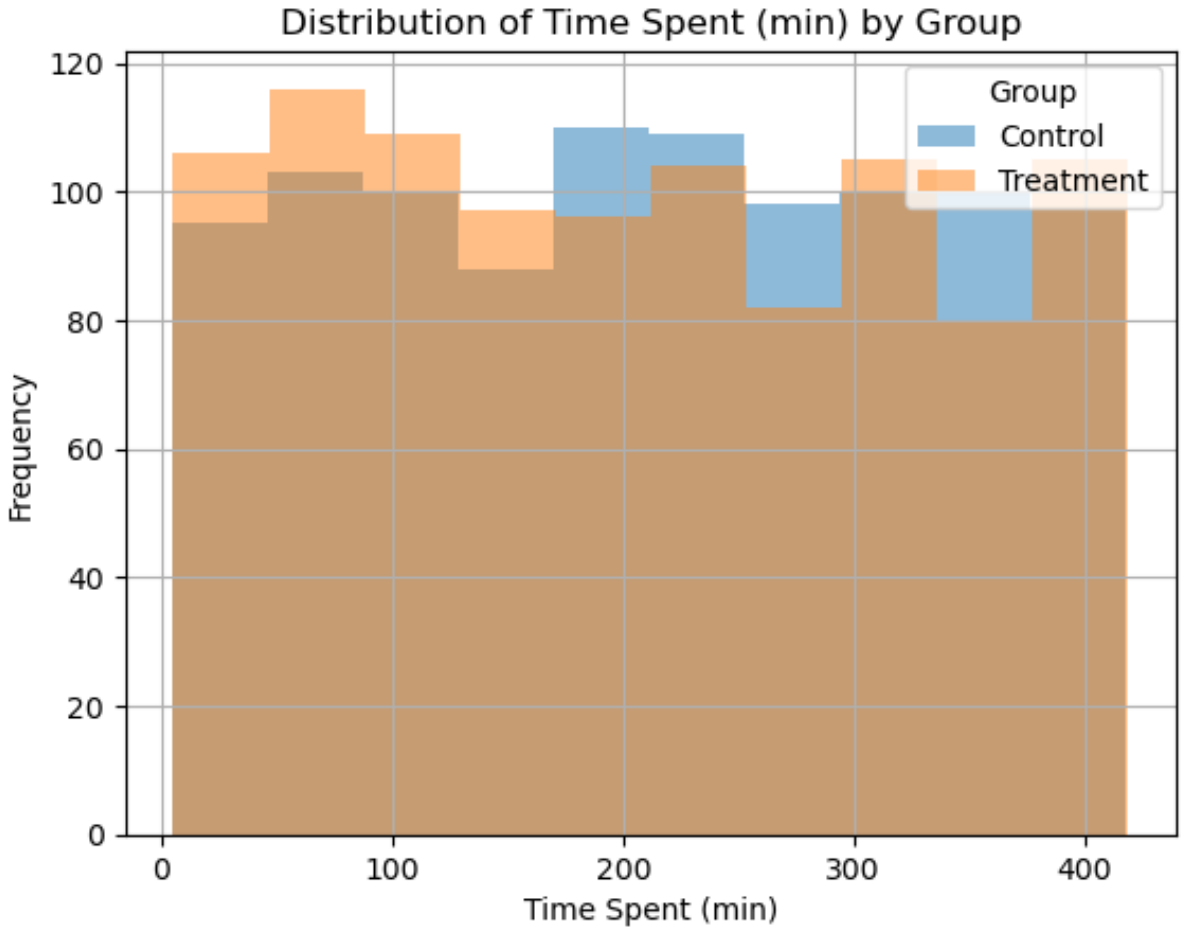
Exploratory Data Analysis

1. Distribution of Key Metrics

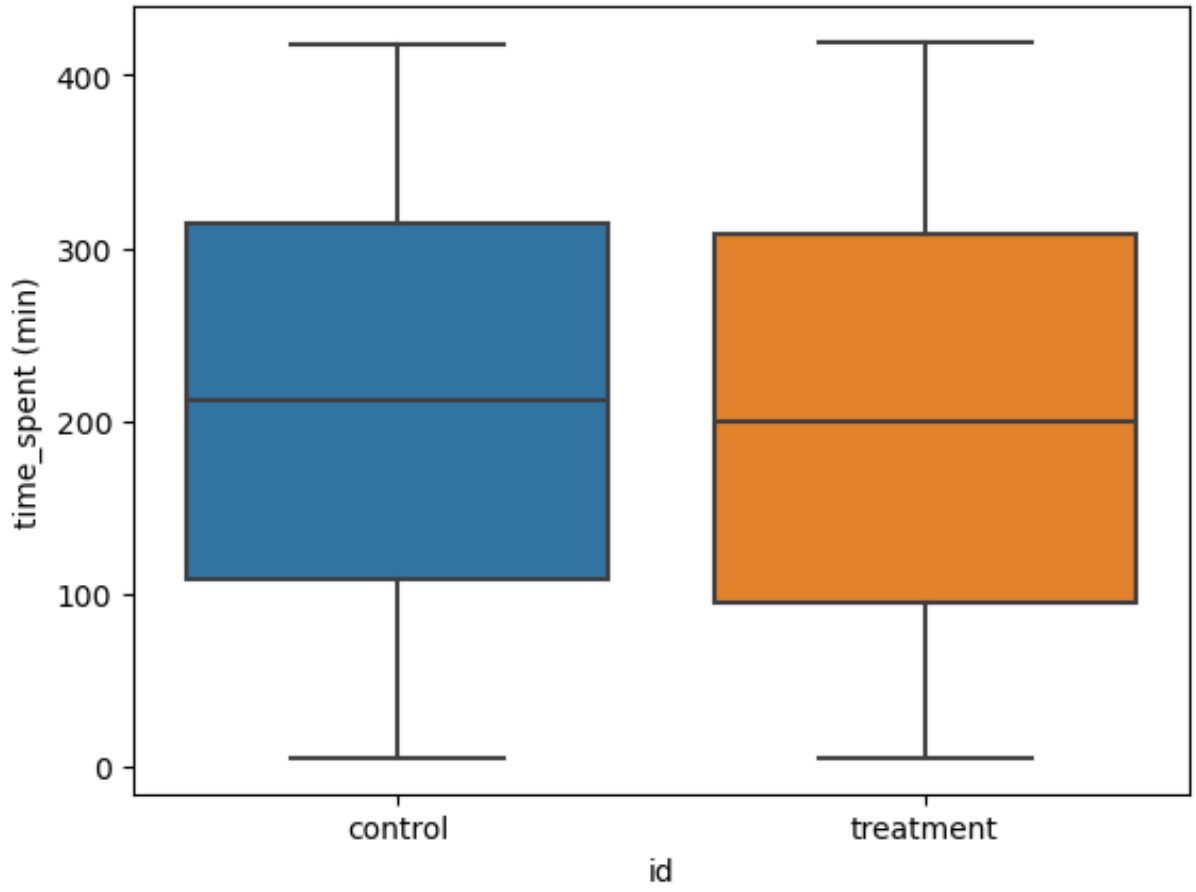
- **Click-Through Rate (CTR)**:
  - Control Group: **65%**
  - Treatment Group: **36%**
  - Observation: The treatment group showed a decrease in CTR compared to the control group.
- **Conversion Rate**:
  - Control Group: **42%**
  - Treatment Group: **13%**
  - Observation: Conversion rates were lower in the treatment group compared to the control group.
- **Time Spent on the Website**:
  - Control Group: Mean = **212 min**, Median = **212.5 min**.
  - Treatment Group: Mean = 205 **min**, Median = 199.0 **min**.
  - Observation: Users in the treatment group spent slightly less time than users in the control group

2. Visualizations

- **Histogram of Time Spent**:

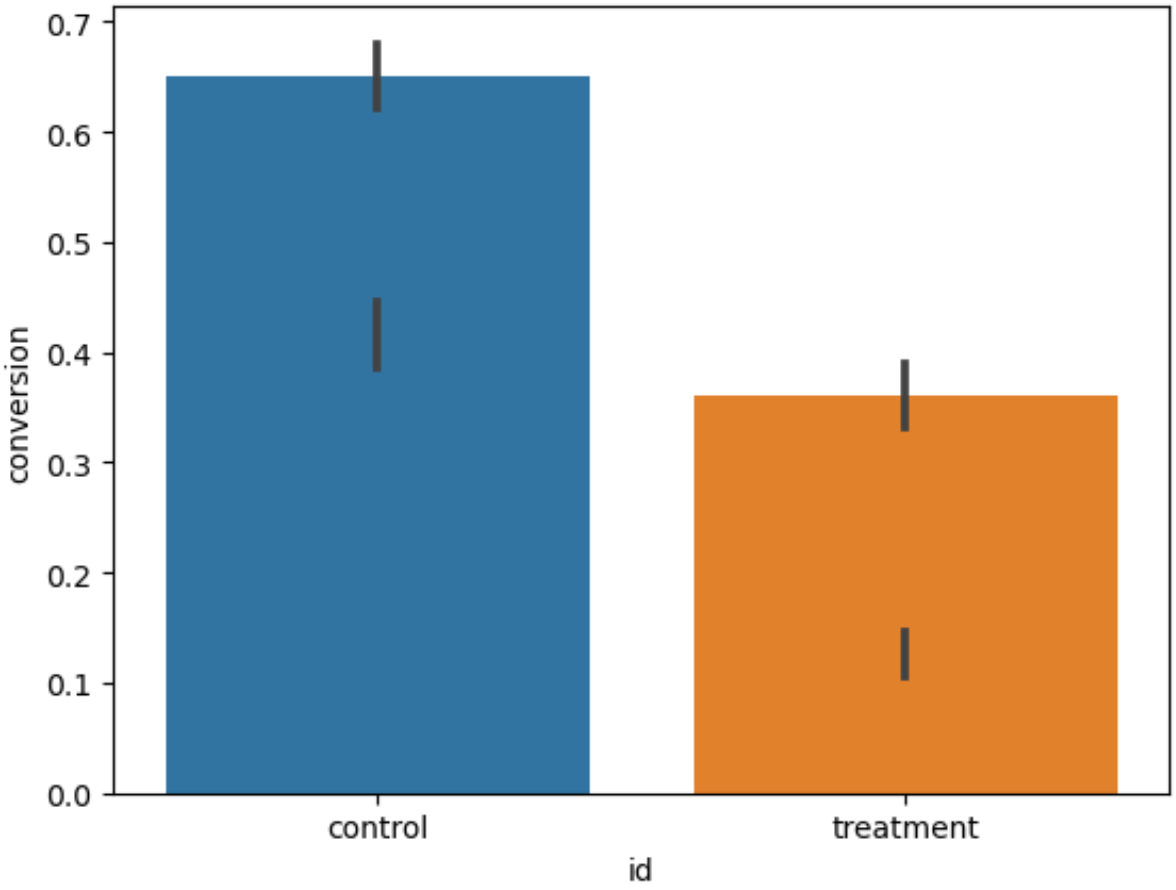


•The histogram revealed that the distribution of time spent on the site for the treatment group was similar to the time spent for the control group



•The boxplot comparison indicated absence of significant outliers in the treatment group.

•**Bar Charts for CTR and Conversion**:



•A bar chart comparing CTR and conversion rates between groups highlighted a significant difference in performance metrics.

3. Correlation Analysis

	time_spent (min)	click	conversion
time_spent (min)	1.000000	0.049953	-0.001210
click	0.049953	1.000000	0.601671
conversion	-0.001210	0.601671	1.000000

- shows that click is strongly positively correlated with conversion with 0.6 correlation coefficient

Hypothesis Testing

1. Click-Through Rate(CTR)

•Null Hypothesis (H_o): There is no difference in CTR between the control and treatment groups.

•Alternative Hypothesis (H_1): The treatment group has a different CTR than the control group.

•**Result**:

- Z-statistic: **12.97**
- P-value: **1.81 x 10^-38**
- Conclusion: With a significance level of **0.05**, we reject the null hypothesis. This indicates that the difference in CTR between the control and treatment groups is significant

2. Conversion Rate(CR)

•Null Hypothesis (H_o): There is no difference in CR between the control and treatment groups.

•Alternative Hypothesis (H_1): The treatment group has a different CR than the control group.

•**Result**:

- Z-statistic: **14.65**
- P-value: **1.39 x 10^-48**
- Conclusion: With a significance level of **0.05**, we reject the null hypothesis. This indicates that the difference in CR between the control and treatment groups is significant

3. Time spent

•Null Hypothesis (H_o): There is no difference in time spent between the control and treatment groups.

•Alternative Hypothesis (H_1): The treatment group has a different CR than the control group.

•Test used: 2-sample t-test

•**Result**:

- T-statistic: **1.30**
- P-value: **0.19**
- Conclusion: With a significance level of **0.05**, we fail to reject the null hypothesis. This indicates that the difference in average time spent is not significant

Effect Size Analysis

Lets quantify the effect sizes

Click-Through Rate (CTR)

- **Absolute Effect Size**: **-0.29**
  - The treatment group's click-through rate was 29 percentage points lower than the control group's.
- **Relative Effect Size**: **-44.55%**
  - This represents a **44.55% decrease** in CTR for the treatment group relative to the control group.
- **Interpretation**: The new design substantially reduced the likelihood of users clicking compared to the old design.

Conversion Rate

- **Absolute Effect Size**: **-0.291**
  - The treatment group's conversion rate was 29.1 percentage points lower than the control group's.
- **Relative Effect Size**: **-69.95%**
  - This corresponds to a **69.95% decrease** in conversion rates for the treatment group relative to the control group.
- **Interpretation**: The new design had a significantly negative impact on user conversions, almost reducing conversions by 70%.

Time Spent on the Website

- **Absolute Effect Size**: **-6.97 minutes**
  - Users in the treatment group spent an average of 6.97 minutes less on the site compared to the control group.
- **Relative Effect Size**: **-3.29%**
  - This reflects a **3.29% decrease** in time spent on the website by the treatment group relative to the control group.
- **Interpretation**: While the decrease in time spent is less dramatic compared to CTR and conversion rates, it still suggests a small negative effect of the new design on user engagement.

Summary of Effect Sizes

- **Click-Through Rate**: The new design led to a **substantial decrease**, with an absolute effect size of **-0.29** and a relative effect size of **-44.55%**.
- **Conversion Rate**: The new design caused a **dramatic reduction**, with an absolute effect size of **-0.291** and a relative effect size of **-69.95%**.
- **Time Spent**: The effect on time spent was less pronounced, with an absolute effect size of **-6.97 minutes** and a relative effect size of **-3.29%**.

Overall Interpretation

The negative effect sizes across all metrics—especially the significant decreases in CTR and conversion rates—indicate that the treatment (new design) performed worse than the control (old design). These results strongly suggest that the new design did not improve user engagement or outcomes and may require re-evaluation or redesign.

Confidence Intervals

- **CTR Difference**: The confidence interval for the difference in CTR was -**0.33, -0.25** indicating that the true difference is likely within this range.
- **Conversion Rate Difference**: The confidence interval for the difference in conversion rates was **-0.327, -0.254.**
- **Time Spent Difference**: was not significant therefore confidence interval wasn't needed

Recommendations

Based on the results of this AB test, the following recommendations are proposed:

1. **Revert to the Control Design**:
   - The treatment design underperformed across all key metrics: click-through rate, conversion rate, and time spent. The control design is currently more effective in engaging users and driving conversions. Reverting to the control design is advised while further improvements are explored.
2. **Identify Issues in the Treatment Design**:
   - Conduct qualitative research (e.g., user feedback, heatmaps, or usability tests) to understand why the new design negatively impacted user behaviour.