

```
In [4]: import pandas as pd

test_data = pd.read_csv(r"C:\Users\pc\Downloads\ab_test_data - ab_test_data.csv (1).csv")
test_data['timestamp'] = pd.to_datetime(test_data['timestamp'])

print('Sample size: a - {}, b - {}'.format(*test_data.groupby('test_group').size()))
print('Conversions: a - {}, b - {}'.format(*test_data[test_data['conversion'] == 1].groupby('test_group').size()))
print('Conversion Rate: a - {}, b - {}'.format(*(test_data.groupby('test_group')['conversion'].mean()*100).round(1)))
print(f'Date Start: {test_data['timestamp'].dt.date.astype(str).min()}')
print(f'Date Finish: {test_data['timestamp'].dt.date.astype(str).max()}')
print(f'Test Duration: {(test_data['timestamp'].dt.date.max() - test_data['timestamp'].dt.date.min()).days}')

Sample size: a - 10013, b - 9985
Conversions: a - 611, b - 889
Conversion Rate: a - 6.1, b - 8.9
Date Start: 2023-07-03
Date Finish: 2023-07-25
Test Duration: 22
```

```
In [5]: from scipy import stats

alpha = 0.05

observed = pd.crosstab(test_data['test_group'].values, test_data['conversion'].values)
statistic, pvalue, dof, expected_values = stats.chi2_contingency(observed)

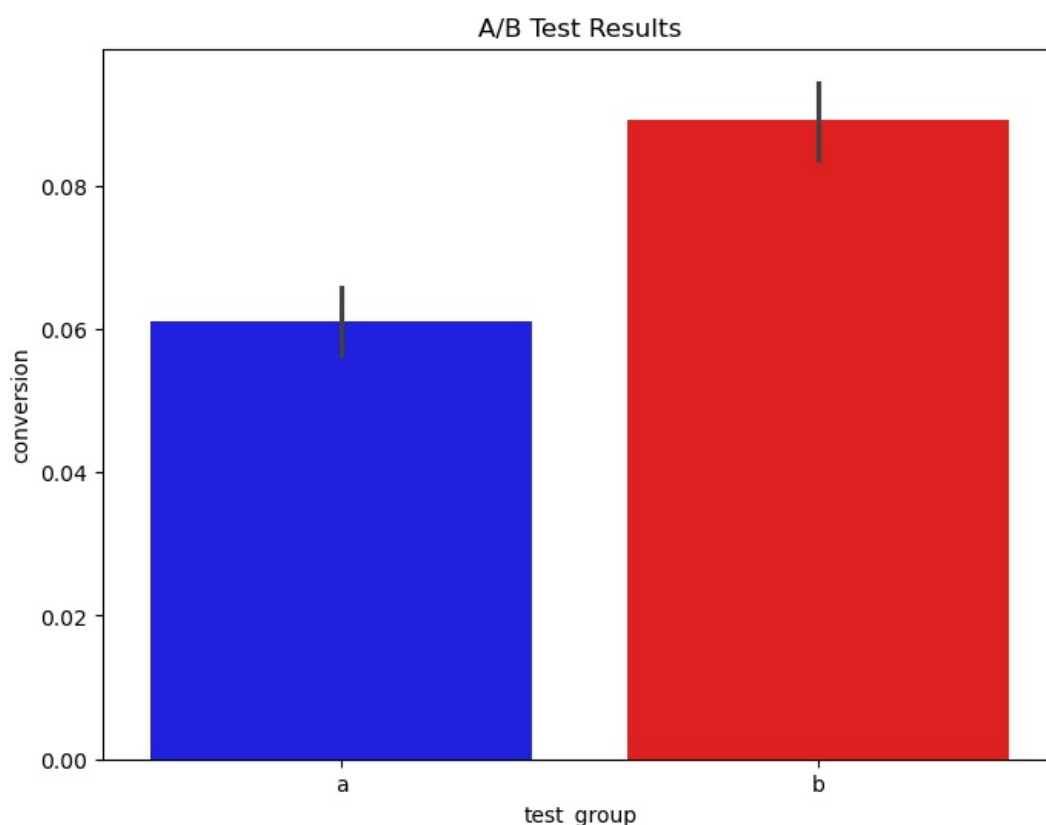
print(f't-statistic: {round(statistic, 2)}, p-value: {round(pvalue, 2)}')

if pvalue < alpha:
    print('The difference is statistically significant, Null Hypothesis is rejected.')
else:
    print('The difference is insignificant, Null Hypothesis cannot be rejected.')

t-statistic: 56.14, p-value: 0.0
The difference is statistically significant, Null Hypothesis is rejected.
```

```
In [15]: import matplotlib.pyplot as plt
import seaborn as sns

plt.figure(figsize=(8, 6))
barplot = sns.barplot(x=test_data['test_group'], y=test_data['conversion'], palette=['blue', 'red'])
plt.title('A/B Test Results')
plt.show()
```



```
In [10]: import matplotlib.pyplot as plt
import seaborn as sns

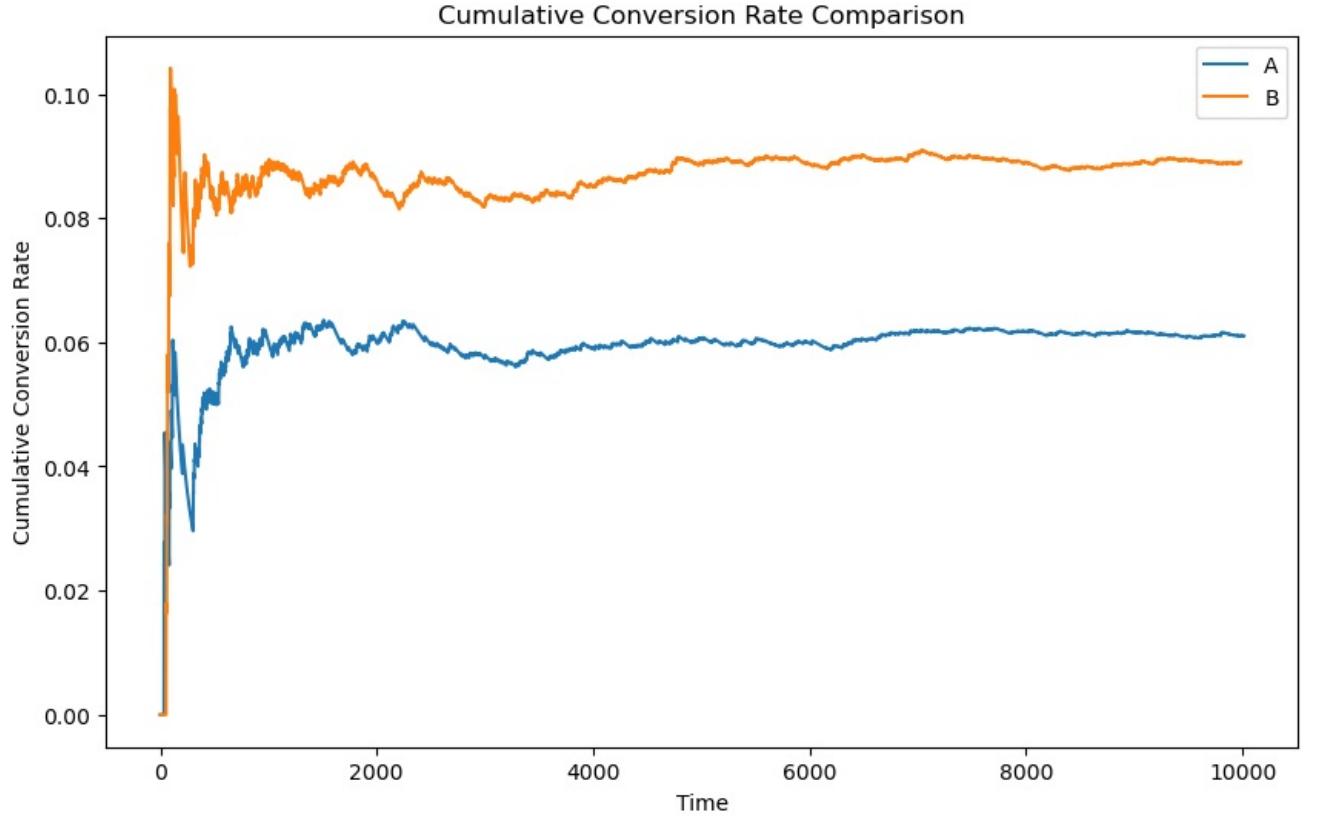
test_data = test_data.sample(frac=1).reset_index(drop=True)

cumulative_metric_a = test_data[test_data['test_group'] == 'a'].sort_values(by='timestamp')['conversion'].expanding().sum().reset_index()
cumulative_metric_b = test_data[test_data['test_group'] == 'b'].sort_values(by='timestamp')['conversion'].expanding().sum().reset_index()

plt.figure(figsize=(10, 6))
plt.plot(cumulative_metric_a, label='A')
plt.plot(cumulative_metric_b, label='B')

plt.title('Cumulative Conversion Rate Comparison')
plt.xlabel('Time')
plt.ylabel('Cumulative Conversion Rate')

plt.legend()
plt.show()
```



In [4]: #Çıkarımlar: A grubunun dönüşüm oranı daha yüksekse, %50 indirimli teklifi sunmanın etkili olduğu sonucuna varılabilir.

A grubunun dönüşüm oranı daha yüksekse, %50 indirimli teklifi sunmanın etkili olduğu sonucuna varılabilir. İstatistiksel test sonuçlarına göre, eğer p-değeri belirli bir eşik değerden düşükse (genellikle 0.05), bu, iki grup arasında anlamlı bir fark olduğunu gösterir. Bu öneriler, A/B testi sonuçlarını değerlendirirken dikkate alınabilecek genel stratejilerdir. Testin spesifik bağlamına ve hedeflerine göre daha detaylı analizler yapılabilir.