Mobile app A|B test analysis

January 8, 2025

[200]:
```
import pandas as pd
from scipy import stats as sc
import seaborn as sb
import matplotlib.pyplot as plt
```

[142]:
```
test_data = pd.DataFrame (data = {'test_group' : ['a']*3490 + ['b']*6897,
                                  'conversion' : [1]*458 + [0]*(3490-458) +␣
 ↪[1]*989 + [0]*(6897-989)})
```

[144]:
```
test_data
```

[144]:

	test_group	conversion
0	a	1
1	a	1
2	a	1
3	a	1
4	a	1
…	…	…
10382	b	0
10383	b	0
10384	b	0
10385	b	0
10386	b	0

[10387 rows x 2 columns]

Visualization of the distribution

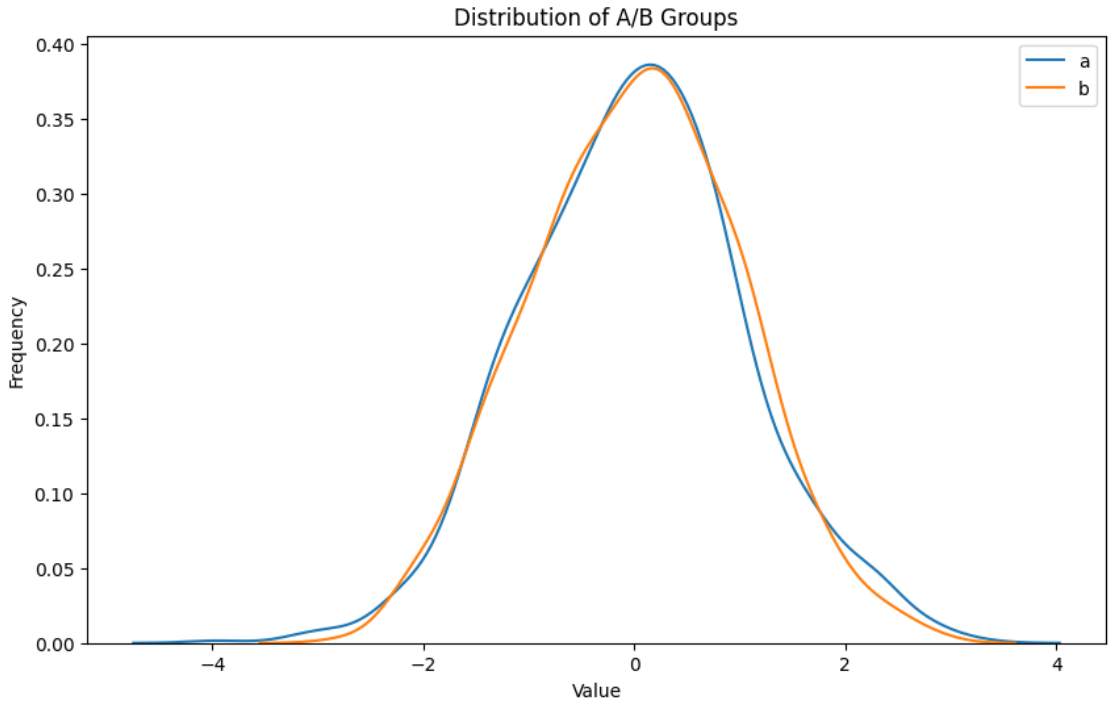
[202]:
```
plt.figure(figsize=(10, 6))

sb.kdeplot(sc.norm.rvs(size=1000))
sb.kdeplot(sc.norm.rvs(size=1000))

plt.title('Distribution of A/B Groups')
plt.xlabel('Value')
plt.ylabel('Frequency')
```

```
plt.legend(['a', 'b'])
plt.show()
```

```
/opt/conda/envs/anaconda-ai-2024.04-py310/lib/python3.10/site-
packages/seaborn/_oldcore.py:1119: FutureWarning: use_inf_as_na option is
deprecated and will be removed in a future version. Convert inf values to NaN
before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
/opt/conda/envs/anaconda-ai-2024.04-py310/lib/python3.10/site-
packages/seaborn/_oldcore.py:1119: FutureWarning: use_inf_as_na option is
deprecated and will be removed in a future version. Convert inf values to NaN
before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
```



Calculating the conversion rate in groups

[232]:
```
conversion_rates = round(test_data.groupby('test_group')['conversion'].mean() *␣
↪100, 2)
```

[234]:
```
conversion_rates
```

[234]:
```
test_group
a    13.12
b    14.34
```

```
Name: conversion, dtype: float64
```

0.0.1 To test the hypotheses and calculate the statistical value and p-value, chosed the statistical criterion t-test from descriptive statistics and t-test_ind.

t-test from descriptive statistics

[146]:
```
test_data.groupby('test_group').describe()
```

[146]:

	conversion							
	count	mean	std	min	25%	50%	75%	max
test_group								
a	3490.0	0.131232	0.337702	0.0	0.0	0.0	0.0	1.0
b	6897.0	0.143396	0.350501	0.0	0.0	0.0	0.0	1.0

[148]:
```
ttest_from_desc_stats = sc.ttest_ind_from_stats(0.337702, 0.131232, 3490, 0.
↪350501, 0.143396, 6897, equal_var=True, alternative='two-sided')
```

[150]:
```
ttest_from_desc_stats
```

[150]:
```
Ttest_indResult(statistic=-4.41900612553385, pvalue=1.001614474742366e-05)
```

[152]:
```
alfa = 0.05
if ttest_from_desc_stats.pvalue < alfa:
    print ('Difference is statistic signifacant by t-test from descriptive␣
↪statistics')
else:  print ('Difference is NOT statistic signifacant by t-test from␣
↪descriptive statistics')
```

```
Difference is statistic signifacant by t-test from descriptive statistics
```

t-test_ind

[154]:
```
ttest_ind = sc.ttest_ind (test_data[test_data['test_group'] ==␣
↪'a']['conversion'],
                   test_data[test_data['test_group'] == 'b']['conversion'],
                    alternative='less')
```

[156]:
```
ttest_ind
```

[156]:
```
TtestResult(statistic=-1.6910814561098335, pvalue=0.04542562618108612,
df=10385.0)
```

[158]:
```
alfa = 0.05
if ttest_ind.pvalue < alfa:
    print ('Difference is statistically significant, Null Hypothesis is␣
↪rejected')
```

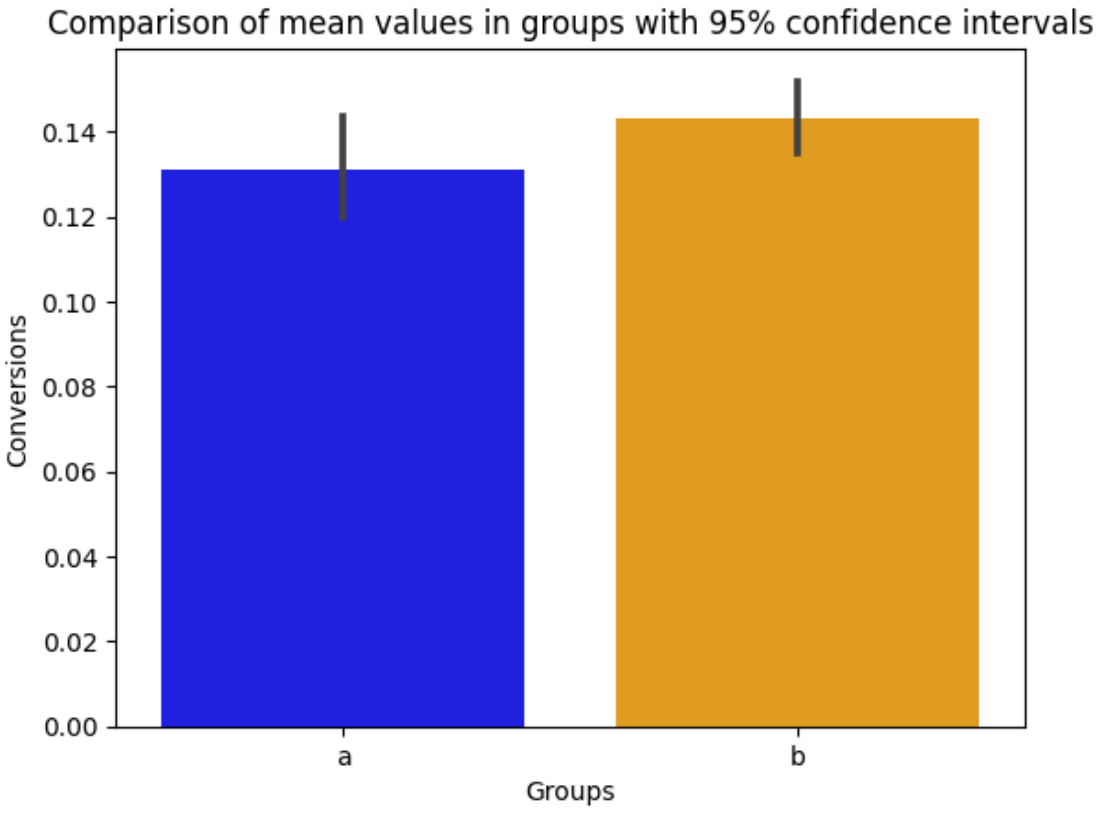
```
else:  print ('Difference is NOT statistic signifacant, Null Hypothesis cannot␣
 ↪rejected')
```

```
Difference is statistically significant, Null Hypothesis is rejected
```

To test the hypotheses, two independent t-tests based on the Student's test were conducted - t-test from descriptive statistics and ttest_ind for the averages of two independent samples, the test results are statistically significant, as follows: t-test from descriptive statistics statistic=-4. 41900612553385, pvalue=1.001614474742366e-05, according to the results of ttest_ind statistic=-1. 6910814561098335, pvalue=0.04542562618108612, df=10385.0, so since pvalue in both cases is less than alpha, this confirms that the test results are statistically significant, we can reject the null hypothesis and accept the alternative one.

0.0.2 Build a visualization to compare mean values in groups with 95% confidence intervals

[160]:
```
sb.barplot(x = test_data['test_group'], y = test_data['conversion'], palette =␣
 ↪['blue', 'orange'])
plt.xlabel('Groups')
plt.ylabel('Conversions')
plt.title('Comparison of mean values in groups with 95% confidence intervals')
plt.show()
```



0.0.3 Building a graph showing the change in conversion over time

Calculate the cumulative average that will be the conversion variable over time

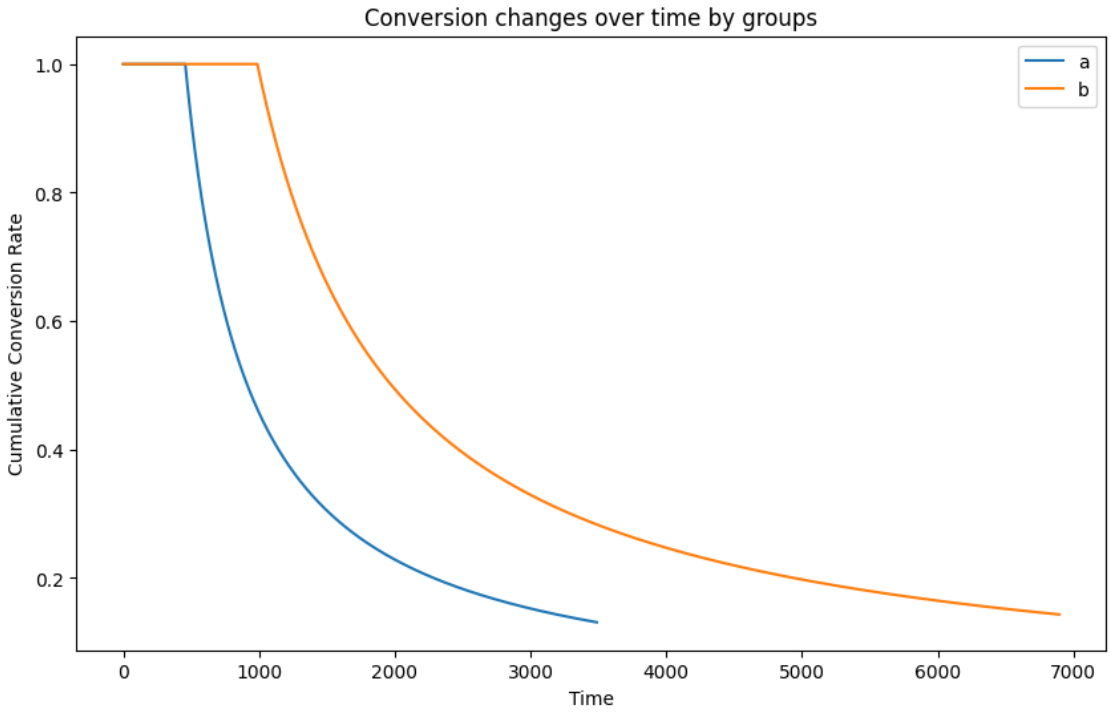
[174]:
```
cumulative_metric_a = test_data[test_data['test_group'] == 'a']['conversion'].
↪expanding().mean().reset_index(drop=True)
cumulative_metric_b = test_data[test_data['test_group'] == 'b']['conversion'].
↪expanding().mean().reset_index(drop=True)
```

[177]:
```
cumulative_metric_a
cumulative_metric_b
```

[177]:
```
0       1.000000
1       1.000000
2       1.000000
3       1.000000
4       1.000000
          …
6892    0.143479
6893    0.143458
6894    0.143437
6895    0.143416
6896    0.143396
Name: conversion, Length: 6897, dtype: float64
```

Building a chart

[181]:
```
plt.figure(figsize=(10, 6))
plt.plot(cumulative_metric_a, label='a')
plt.plot(cumulative_metric_b, label='b')
plt.xlabel('Time')
plt.ylabel('Cumulative  onversion Rate')
plt.title('Conversion changes over time by groups')
plt.legend()
plt.show()
```



Conclusions: Based on the results of the independent t-tests, we can conclude that the test results are statistically significant, and therefore we can reject the null hypothesis and accept the alternative.

According to the alternative hypothesis, the product uses an alternative design of the subscription screen, offering a weekly subscription of premium features at the same price of $4.99, but with a 50% discount. It is expected that the implementation of this design will lead to an increase in conversion from installation to payment, which will have a positive impact on business performance. Group 'B' (with the alternative design of the subscription screen) showed a higher conversion rate (14.36%) compared to group 'A' (13.12%).

This shows that the new design is more effective in converting users. The results of the experiment have sufficient statistical power for decision-making because the effect of the difference between the groups is real, not random. The cumulative conversion of group 'B' consistently exceeds that of group 'A' This confirms the reliability of the results, as the effect persists over time.

[ ]: