A/B testing for New Website Redesign

Use and Explanation of Fisher's Exact Test

By Matt Curcio

A. Executive Summary

This report discusses the current redesign of the landing page to determine if conversion rate is increased.

1. The total number of observations is **n = 20,123** for this experiment.
    - The Conversion rate when exposed to the treatment was 11.44%.
    - The Conversion rate when NOT exposed to the treatment page was 10.77%

Conclusion:

1. The new landing page redesign did not have any statistically significant change over the original landing page.
2. I recommend that experimentation be continued to find a higher conversion rate.
3. The Fisher's Exact Test is used to calculate the probability and the odds for this to occur.
    - **The Probability of the two events being equivalent are < 2.2e-16**.
    - **The Odds are on average 0.0156601 to 1 for this event to occur**.

B. Discussion

2 x 2 Contingency Table for Treatment Versus New Images

	Treatment-Yes	Treatment-No	Marginal Sums
New_images-Yes	1151	8982	10133
New_images-No	8906	1084	9990
Sums	10057	10066	20123

1. **Treatment**

- 49.9776375 % of the 20,123 observations viewed the treatment page.
- The conversion rate of the Treatment page is 11.4447648 %.

2. **NOT-Treatment**

- 50.0223625 % of the 20,123 observations did NOT view the treatment page
- The conversion rate of the Original page is 10.7689251 %.

C. Results

- Load Redesign Test Data

```
library(readr)
redesign <- read_csv("redesign.csv",
     col_types = cols(treatment = col_factor(levels = c("yes","no")),
                      new_images = col_factor(levels = c("yes","no")),
                      converted = col_factor(levels = c("0","1"))))
# View(redesign)
```

1. The summary below shows that there no missing values.
2. Redesign test data
    - “treatment” - The set of this categorical variable are {“yes”,“no”}
    - “new_images” - The set of this categorical variable are {“yes”,“no”}
    - “converted” - This discrete output variable is {“0”,“1”}.
        - “0” is NOT converted, “1” IS converted.

```
summary(redesign)
```

```
##  treatment   new_images  converted
##  yes:20242   yes:20242   0:35895
##  no :20242   no :20242   1: 4589
```

3. Filter dataset based on 2 x 2 input variables dsiplaying the number of conversions.

```
redesign_TT = subset(redesign, treatment == "yes" & new_images == "yes" &
                     converted == "1")
#nrow(redesign_TT)

redesign_TF = subset(redesign, treatment == "yes" & new_images == "no" &
                     converted == "0")
#nrow(redesign_TF)

redesign_FT = subset(redesign, treatment == "no" & new_images == "yes" &
                     converted == "0")
#nrow(redesign_FT)

redesign_FF = subset(redesign, treatment == "no" & new_images == "no" &
                     converted == "1")
#nrow(redesign_FF)
```

```
- 1151 observations for Treatment=Yes, New Images=Yes, and Converted=Yes.

- 8906 observations for Treatment=Yes, New Images=No, and Converted=NO

- 8982 observations for Treatment=No, New Images=Yes, and Converted=Yes.

- 1084 observations for Treatment=No, New Images=No, and Converted=NO
```

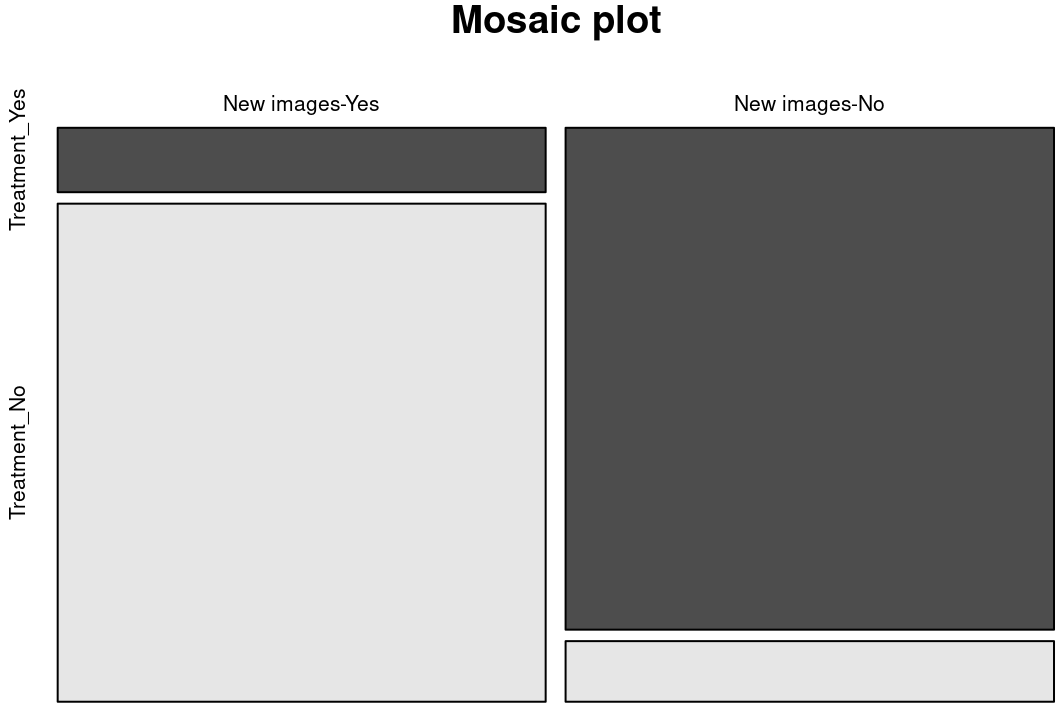
4. Conversion Ratio
    - The conversion rate of the Treatment page is 11.4447648 %.
    - The conversion rate of the Original page is 10.7689251 %.
5. Create 2x2 Contingency table

```
Contingency_table <- data.frame("Treatment_Yes" = c(1151, 8982),
                                "Treatment_No" = c(8906, 1084),
                                row.names = c("New images-Yes", "New images-No"),
                                stringsAsFactors = FALSE)

colnames(Contingency_table) <- c("Treatment_Yes", "Treatment_No")

Contingency_table
```

```
##                Treatment_Yes Treatment_No
## New images-Yes          1151         8906
## New images-No           8982         1084
```

6. Mosaic plot of Conversion testing

```
mosaicplot(Contingency_table,
           main = "Mosaic plot",
           color = TRUE
           )
```



7. Fisher Exact Test

```
f_test <- fisher.test(Contingency_table)
f_test
```

```
##
##  Fisher's Exact Test for Count Data
##
## data:  Contingency_table
## p-value < 2.2e-16
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.01424845 0.01707176
## sample estimates:
## odds ratio
## 0.01561694
```

8. P-value for Fisher's test
    - From the Fisher's Exact Test output and specifically from the test$p.value we see that the p-value (Pr < 2.2e-16) is much less than the significance level of 1% for a two-tailed test.
    - ODDS??

D. Conclusion and Interpretation

From the Fisher's Exact Test output and specifically from the test$p.value we see that the p-value (< 2.2e-16) is much less than the significance level of 1% for a two-tailed test.

In statistical testing, if the p-value is less than the significance level, **we can reject the null hypothesis**.

In this circumstance, rejecting the null hypothesis for the Fisher's exact test of independence means that there is a NO significant relationship between the two categorical variables (Treatments and New Images).