A/B testing

Mingwei Wu

Introduction for A/B testing

A general framework of hypothesis testing between two groups to establish a casual relationship between actions and results.

Understand problem&objectives

Come up with hypothesis

Code change and testing

Run experiment and monitor

Data analysis

Blackground

Website clicking is the key point to the web. Assuming to check the cliked like in the website, track the metrics and run the experiment.

Data Description

visit_date: the date of customer visit the website.

clicked_adopt_today: customer click at visit day.

condition: different group

time_spent_homepage_sec: customer spend a time of second in the homepage

clicked_article: article for click by customer

clicked_like: take like for article

clicked_share: clicked share

```
##   visit_date clicked_adopt_today
## 1 2017-01-01                   1
## 2 2017-01-02                   1
## 3 2017-01-03                   0
## 4 2017-01-04                   1
## 5 2017-01-05                   1
## 6 2017-01-06                   0
```

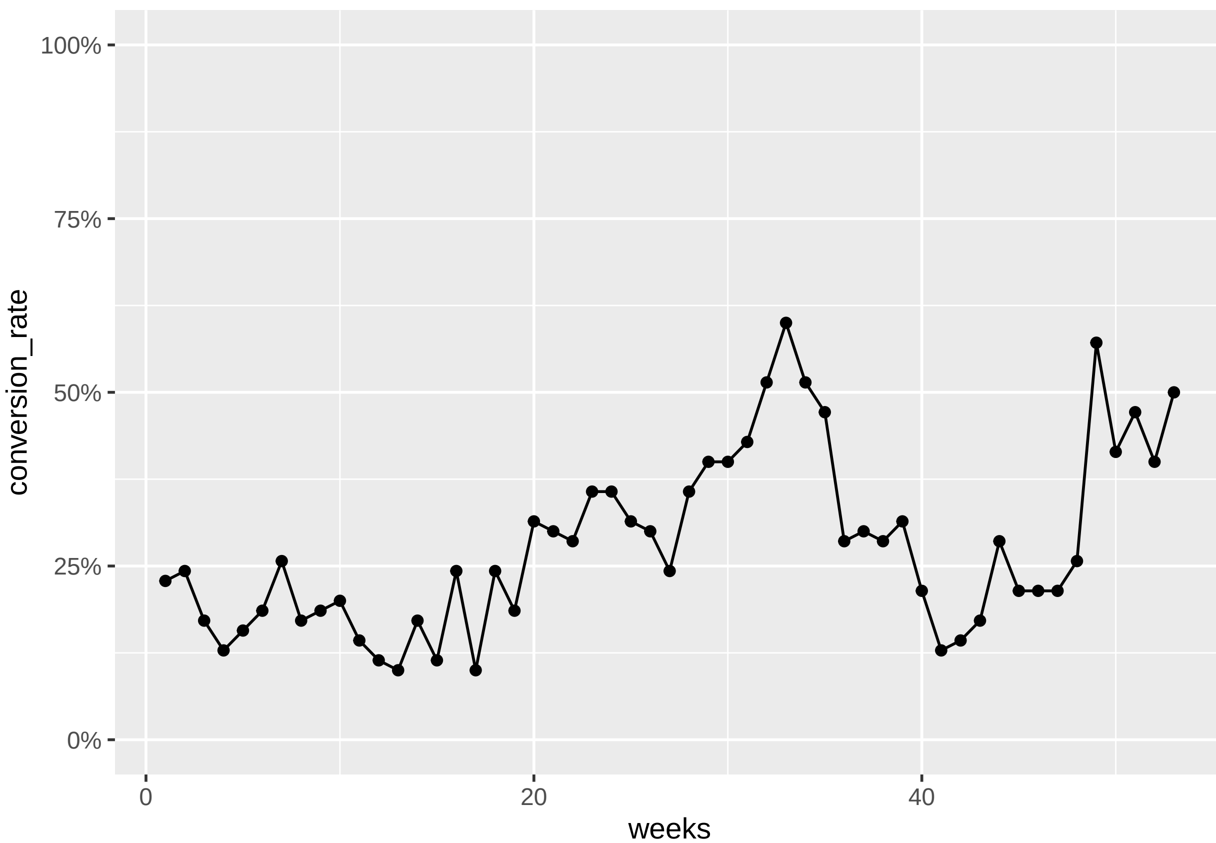
```
## -- Attaching packages --------------------------------- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.2     v purrr   0.3.4
## v tibble  3.0.4     v dplyr   1.0.2
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0

## -- Conflicts ------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

##
## Attaching package: 'lubridate'

## The following object is masked from 'package:base':
##
##     date

## [1] "2017-01-01"

## [1] "2017-12-31"

## `summarise()` ungrouping output (override with `.groups` argument)

## # A tibble: 7 x 2
##   `wday(visit_date)` conversion_rate
##                <dbl>           <dbl>
## 1                  1           0.3
## 2                  2           0.277
## 3                  3           0.271
## 4                  4           0.298
## 5                  5           0.271
## 6                  6           0.267
## 7                  7           0.256

## `summarise()` ungrouping output (override with `.groups` argument)

## # A tibble: 6 x 2
##   `week(visit_date)` conversion_rate
##                <dbl>           <dbl>
## 1                  1           0.229
## 2                  2           0.243
## 3                  3           0.171
## 4                  4           0.129
## 5                  5           0.157
## 6                  6           0.186
```

plot the clciked rate in weeks in 2017. Throught the graph, assume the summer and winter the web cliking is higher than other weeks. provide the hypothesis for the children vacation, is it raise the website clicking?

```
##
## Attaching package: 'scales'
```

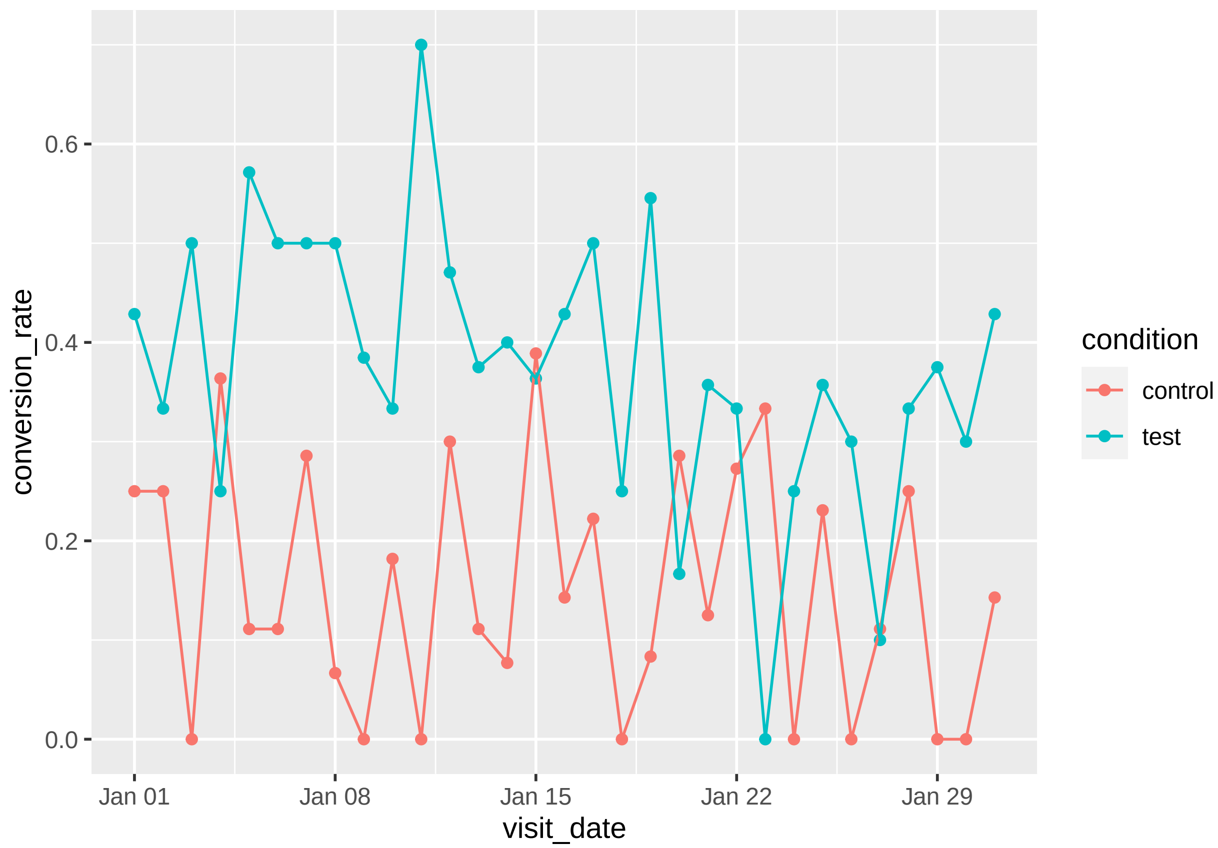
```
## The following object is masked from 'package:purrr':
##
##     discard

## The following object is masked from 'package:readr':
##
##     col_factor
```



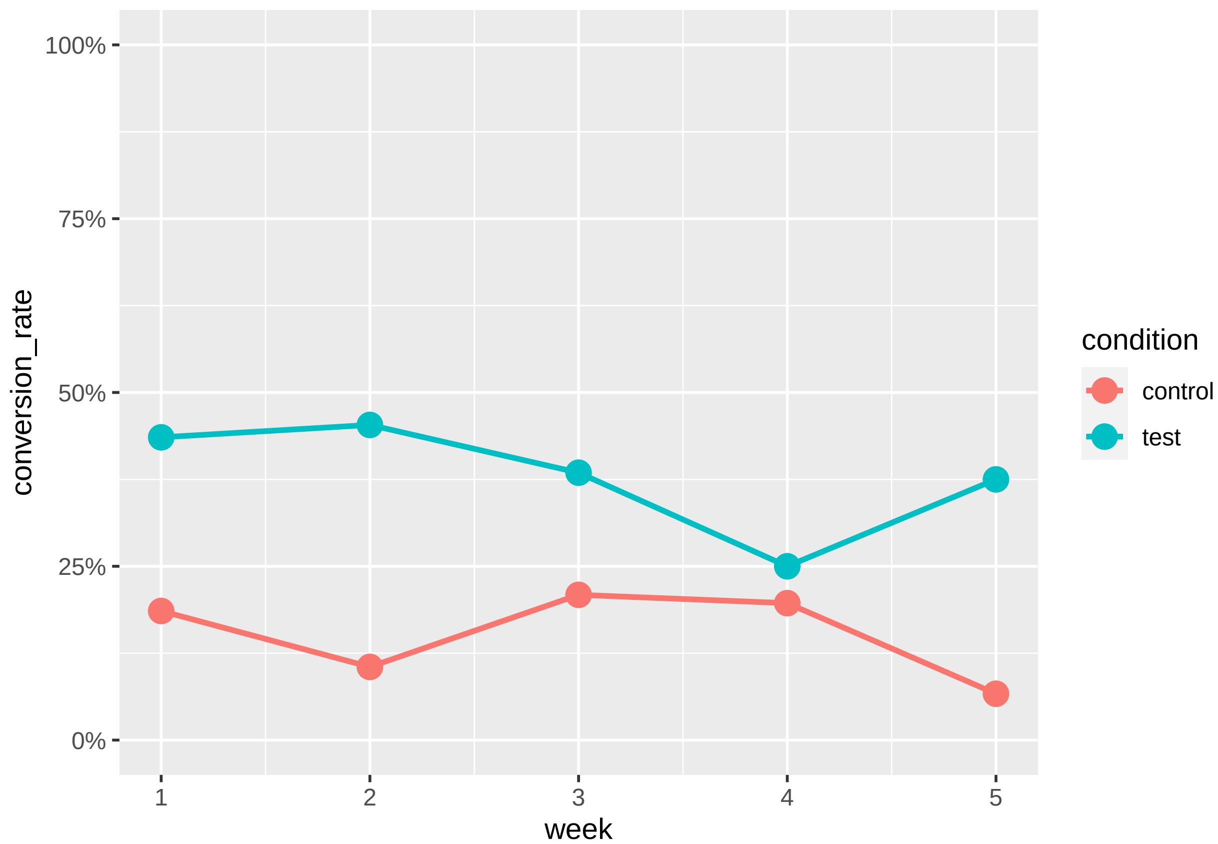

```
## [1] 758

##   visit_date condition clicked_adopt_today
## 1 2018-01-01   control                   0
## 2 2018-01-01   control                   1
## 3 2018-01-01   control                   0
## 4 2018-01-01   control                   0
## 5 2018-01-01      test                   0
## 6 2018-01-01      test                   0
```

After set up 2 group, the testing is runs good in conversion rate than the control group.

```
## `summarise()` regrouping output by 'visit_date' (override with `.groups` argument)
```




```
## # A tibble: 2 x 5
##   term            estimate std.error statistic  p.value
##   <chr>              <dbl>     <dbl>     <dbl>    <dbl>
## 1 (Intercept)        -1.61     0.156    -10.3  8.28e-25
## 2 conditiontest       1.14     0.197      5.77 7.73e- 9

## [1] 194

## `summarise()` regrouping output by 'week' (override with `.groups` argument)
```




```
##   visit_date condition time_spent_homepage_sec clicked_article
## 1 2018-04-01      tips                49.01161               1
## 2 2018-04-01      tips                48.86452               1
## 3 2018-04-01      tips                49.07467               1
## 4 2018-04-01      tips                49.26011               0
## 5 2018-04-01      tips                50.37190               0
## 6 2018-04-01      tips                49.08458               1
##   clicked_like clicked_share
## 1            0             1
## 2            0             0
## 3            0             0
## 4            1             0
## 5            1             0
## 6            0             0
```

```
## # A tibble: 2 x 5
##   term           estimate std.error statistic   p.value
##   <chr>             <dbl>     <dbl>     <dbl>     <dbl>
## 1 (Intercept)      -1.61     0.0219     -73.5 0.
## 2 conditiontools   -0.989    0.0390     -25.4 4.13e-142
```

```
## 
##  Welch Two Sample t-test
## 
## data:  time_spent_homepage_sec by condition
```

```
## t = 0.36288, df = 29997, p-value = 0.7167
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.01850573  0.02691480
## sample estimates:
##  mean in group tips mean in group tools
##            49.99909            49.99489
```

Sequential analysis sample size

```
## One-sided group sequential design with
## 80 % power and 5 % Type I Error.
##            Sample
##             Size
##   Analysis Ratio*  Z   Nominal p  Spend
##          1  0.394 1.99    0.0232 0.0232
##          2  0.789 1.99    0.0232 0.0155
##          3  1.183 1.99    0.0232 0.0113
##      Total                       0.0500
##
## ++ alpha spending:
##  Pocock boundary.
## * Sample size ratio compared to fixed design with no interim
##
## Boundary crossing probabilities and expected sample size
## assume any cross stops the trial
##
## Upper boundary (power or Type I Error)
##           Analysis
##    Theta      1      2      3 Total   E{N}
##   0.0000 0.0232 0.0155 0.0113  0.05 1.1591
##   2.4865 0.3334 0.2875 0.1791  0.80 0.8070
```

```
## [1] 250 500 750
```