Funnel analysis

Tianpei Qian

2023-01-18

TL;DR

- There is a bug or a bad product change starting from March that affects the search page on mobile and the payment page on web.
- Mobile has much higher conversion rate but fewer customers than web. We need to understand whether it is a marketing problem or a product problem. If it is a marketing problem (i.e. mobile attracts customers with higher intentions to buy), then we should try to find more such customers. If it is a product problem (i.e. mobile experience is better than web), then we need to improve the web version.
- The drop-off rate from search to payment and payment to confirmation seems very low. We probably have a bad search engine, a poor product catalog, a substandard product page or an inconvenient payment system. We need to do a deep dive to understand the exact reasons.

Step 1: read data

You may use the read_lines function to peek the files.

```
user_table_file <- 'user_table.csv'
home_page_file <- 'home_page_table.csv'
search_page_file <- 'search_page_table.csv'
payment_page_file <- 'payment_page_table.csv'
payment_confirmation_file <- 'payment_confirmation_table.csv'

user <-
  read_csv(
    user_table_file,
    col_types = list(col_integer(), col_date(), col_character(), col_character())
  )

home_page <-
  read_csv(
    home_page_file,
    col_types = list(col_integer(), col_character())
  )

search_page <-
  read_csv(
    search_page_file,
    col_types = list(col_integer(), col_character())
  )

payment_page <-
  read_csv(
    payment_page_file,
```

```
    col_types = list(col_integer(), col_character())
  )

confirmation_page <-
  read_csv(
    payment_confirmation_file,
    col_types = list(col_integer(), col_character())
  )
```

Step 2: data cleaning

Join tables, process factors, etc.

```
funnel <-
  user %>%
  mutate(
    device = factor(device, levels = c("Desktop", "Mobile")),
    sex = factor(sex, levels = c("Male", "Female"))
  ) %>%
  left_join(home_page, by = "user_id") %>%
  mutate(home_page_flag = !is.na(page)) %>%
  select(-page) %>%
  left_join(search_page, by = "user_id") %>%
  mutate(search_page_flag = !is.na(page)) %>%
  select(-page) %>%
  left_join(payment_page, by = "user_id") %>%
  mutate(payment_page_flag = !is.na(page)) %>%
  select(-page) %>%
  left_join(confirmation_page, by = "user_id") %>%
  mutate(confirmation_page_flag = !is.na(page)) %>%
  select(-page)
```

Step 3: exploratory analysis

```
funnel %>%
  summarize(
    n = n(),
    overall = sum(confirmation_page_flag) / sum(home_page_flag),
    home_search = sum(search_page_flag) / sum(home_page_flag),
    search_payment = sum(payment_page_flag) / sum(search_page_flag),
    payment_confirmation = sum(confirmation_page_flag) / sum(payment_page_flag)
  )
```

```
## # A tibble: 1 x 5
##       n overall home_search search_payment payment_confirmation
##   <int>   <dbl>       <dbl>          <dbl>                <dbl>
## 1 90400   0.005         0.5          0.133               0.0750
```

The overall conversion rate is only 0.5%. The biggest drop off occurs from payment to confirmation - only 7% make through.

```
funnel %>%
  group_by(date) %>%
  summarize(
    n = n(),
```

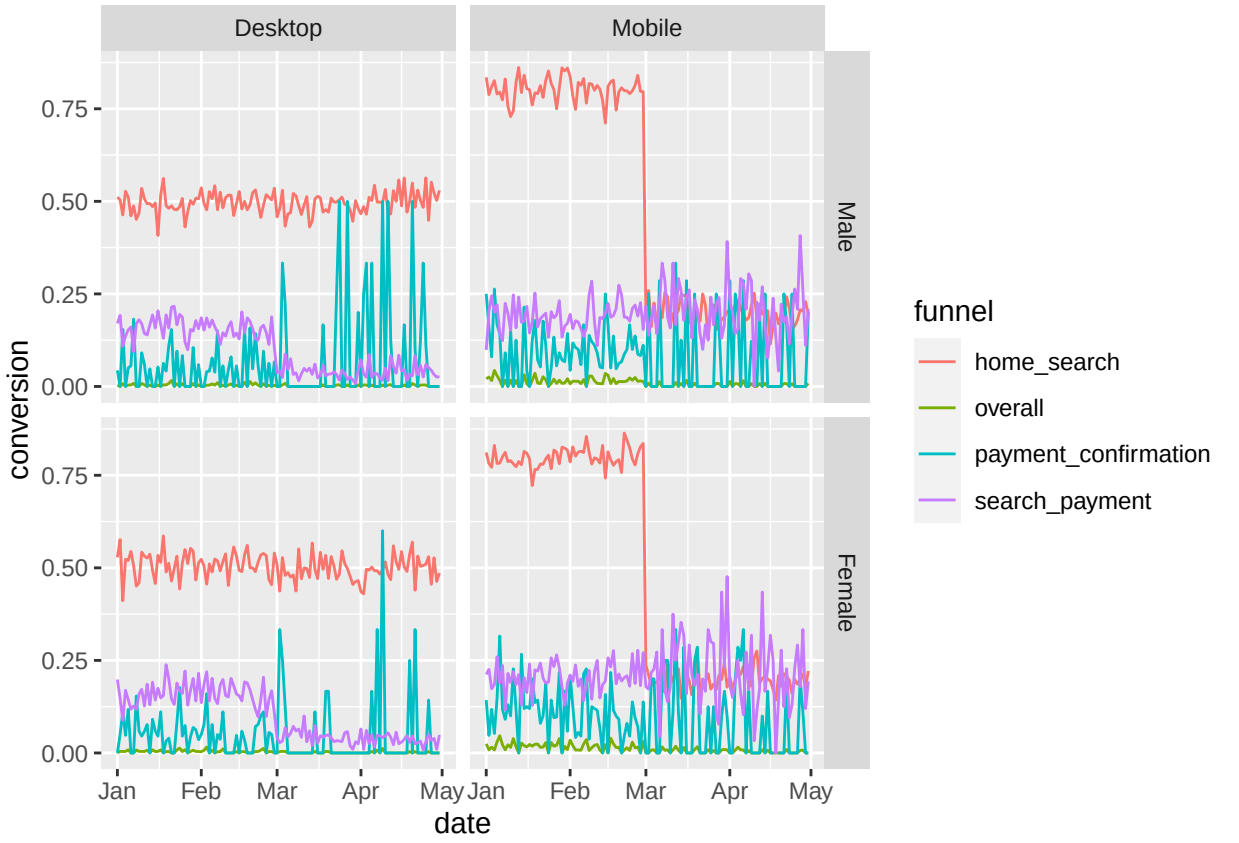
```
    overall = sum(confirmation_page_flag) / sum(home_page_flag),
    home_search = sum(search_page_flag) / sum(home_page_flag),
    search_payment = sum(payment_page_flag) / sum(search_page_flag),
    payment_confirmation = sum(confirmation_page_flag) / sum(payment_page_flag)
  ) %>%
  pivot_longer(overall:payment_confirmation, names_to = "funnel", values_to = "conversion") %>%
  ggplot(aes(date, conversion)) +
  geom_line(aes(color = funnel))
```

There is an obvious drop in conversion from home to search and from search to payment starting from March. Let's do a deeper dive.

```
funnel %>%
  group_by(date, device, sex) %>%
  summarize(
    n = n(),
    overall = sum(confirmation_page_flag) / sum(home_page_flag),
    home_search = sum(search_page_flag) / sum(home_page_flag),
    search_payment = sum(payment_page_flag) / sum(search_page_flag),
    payment_confirmation = sum(confirmation_page_flag) / sum(payment_page_flag)
  ) %>%
  pivot_longer(overall:payment_confirmation, names_to = "funnel", values_to = "conversion") %>%
  ggplot(aes(date, conversion)) +
  geom_line(aes(color = funnel)) +
  facet_grid(sex ~ device)
```

```
## `summarise()` has grouped output by 'date', 'device'. You can override using
```

```
## the `.groups` argument.
```



Since March, the search page on mobile and the payment page on web start to have some problems. Gender doesn't seem to matter much for conversion.

```
funnel %>%
  filter(date < ymd('2015-03-01')) %>%
  group_by(device) %>%
  summarize(
    n = n(),
    overall = sum(confirmation_page_flag) / sum(home_page_flag),
    home_search = sum(search_page_flag) / sum(home_page_flag),
    search_payment = sum(payment_page_flag) / sum(search_page_flag),
    payment_confirmation = sum(confirmation_page_flag) / sum(payment_page_flag)
  )
```

```
## # A tibble: 2 x 6
##   device      n overall home_search search_payment payment_confirmation
##   <fct>   <int>   <dbl>       <dbl>          <dbl>                <dbl>
## 1 Desktop 30100 0.00379       0.504          0.159               0.0473
## 2 Mobile  15100 0.0164        0.8            0.198               0.104
```

Before the problem occurs, the biggest drop off occurs from payment to confirmation, followed by search to payment. They both have big room for improvement. We have a good percentage of customers going from the home page to the search page. Their behavior indicates that they do have some buying intentions. However, most of them bounce after searching. The drop from search to payment means either we have a bad search engine or we just don't have enough good products. The drop from payment to confirmation indicates we probably don't have a good product page to convince the customers to buy (e.g. lack of product

reviews) or the payment system is hard to use.

Mobile performs much better than desktop but attracts fewer customers. We need a deep dive to understand why. If it is a marketing problem (i.e. mobile attracts customers with higher intentions to buy), then we should try to find more such customers. If it is a product problem (i.e. mobile experience is better than web), then we need to improve the web version.