# Free-to-Paid Conversion Analysis

Guy Mor

2025-05-04

```
cat(query, sep = "\n")
```

```
## SELECT
##     ROUND(SUM(first_date_purchased >= first_date_watched) * 100.0 / COUNT(*),
##             2) AS conversion_rate,
##     ROUND(SUM(date_diff_reg_watch) * 1.0 / COUNT(*),
##             2) AS av_reg_watch,
##     ROUND(SUM(date_diff_watch_purch) * 1.0 / COUNT(date_diff_watch_purch),
##             2) AS av_watch_purch
## FROM
##     (SELECT
##         si.student_id,
##             si.date_registered,
##             MIN(se.date_watched) AS first_date_watched,
##             MIN(sp.date_purchased) AS first_date_purchased,
##             ABS(DATEDIFF(MIN(se.date_watched), si.date_registered)) AS date_diff_reg_watch,
##             ABS(DATEDIFF(MIN(sp.date_purchased), MIN(se.date_watched))) AS date_diff_watch_purch
##     FROM
##         student_info si
##     JOIN student_engagement se ON si.student_id = se.student_id
##     LEFT JOIN student_purchases sp ON si.student_id = sp.student_id
##     GROUP BY si.student_id , si.date_registered
##     HAVING first_date_purchased IS NULL
##         OR first_date_watched <= first_date_purchased) AS sub;
```

```
result <- data.frame(
  conversion_rate = 11.29,
  av_reg_watch = 3.42,
  av_watch_purch = 26.25
)

knitr::kable(result, caption = "Summary of Student Engagement Metrics")
```

Table 1: Summary of Student Engagement Metrics

| conversion_rate | av_reg_watch | av_watch_purch |
|---|---|---|
| 11.29 | 3.42 | 26.25 |

## Interpretation

- **Conversion Rate (11.29%)**:
  Slightly more than 1 in 10 students who engaged with a lecture chose to purchase a subscription.
- **Average Registration-to-Engagement Time (3.42 days)**:
  Students typically start watching lectures within just a few days after signing up, which may indicate that the platform is easy to use and effectively encourages new users to get started quickly.
- **Average Engagement-to-Purchase Time (26.25 days)**:
  On average, students wait nearly a month after watching their first lecture before subscribing. This may reflect hesitation or a tendency to wait for special deals before making a purchase.

```
df <- read.csv("data/time_differences.csv")
head(df, 5)
```

```
##   student_id date_registered first_date_watched first_date_purchased
## 1     255193      2021-12-01         2021-12-01                 NULL
## 2     255194      2021-12-01         2021-12-01                 NULL
## 3     255198      2021-12-01         2022-02-17                 NULL
## 4     255199      2021-12-01         2021-12-01                 NULL
## 5     255200      2021-12-01         2021-12-01                 NULL
##   date_diff_reg_watch date_diff_watch_purch
## 1                   0                  NULL
## 2                   0                  NULL
## 3                  78                  NULL
## 4                   0                  NULL
## 5                   0                  NULL
```

```
get_mode <- function(v) {
  v <- na.omit(v)
  uniqv <- unique(v)
  uniqv[which.max(tabulate(match(v, uniqv)))]
}
```

```
df$date_diff_reg_watch <- as.numeric(df$date_diff_reg_watch)
df$date_diff_watch_purch <- suppressWarnings(as.numeric(df$date_diff_watch_purch))


# Registration to Watch
mean_reg_watch   <- mean(df$date_diff_reg_watch, na.rm = TRUE)
median_reg_watch <- median(df$date_diff_reg_watch, na.rm = TRUE)
mode_reg_watch   <- get_mode(df$date_diff_reg_watch)

# Watch to Purchase
mean_watch_purch   <- mean(df$date_diff_watch_purch, na.rm = TRUE)
median_watch_purch <- median(df$date_diff_watch_purch, na.rm = TRUE)
mode_watch_purch   <- get_mode(df$date_diff_watch_purch)

cat("Registration → Watch:\n")
```

```
## Registration → Watch:
```

```r
cat("Mean:   ", round(mean_reg_watch, 2), "\n")
```

```
## Mean:    3.42
```

```r
cat("Median: ", median_reg_watch, "\n")
```

```
## Median:  0
```

```r
cat("Mode:   ", mode_reg_watch, "\n\n")
```

```
## Mode:    0
```

```r
cat("Watch → Purchase:\n")
```

```
## Watch → Purchase:
```

```r
cat("Mean:   ", round(mean_watch_purch, 2), "\n")
```

```
## Mean:    26.25
```

```r
cat("Median: ", median_watch_purch, "\n")
```

```
## Median:  1
```

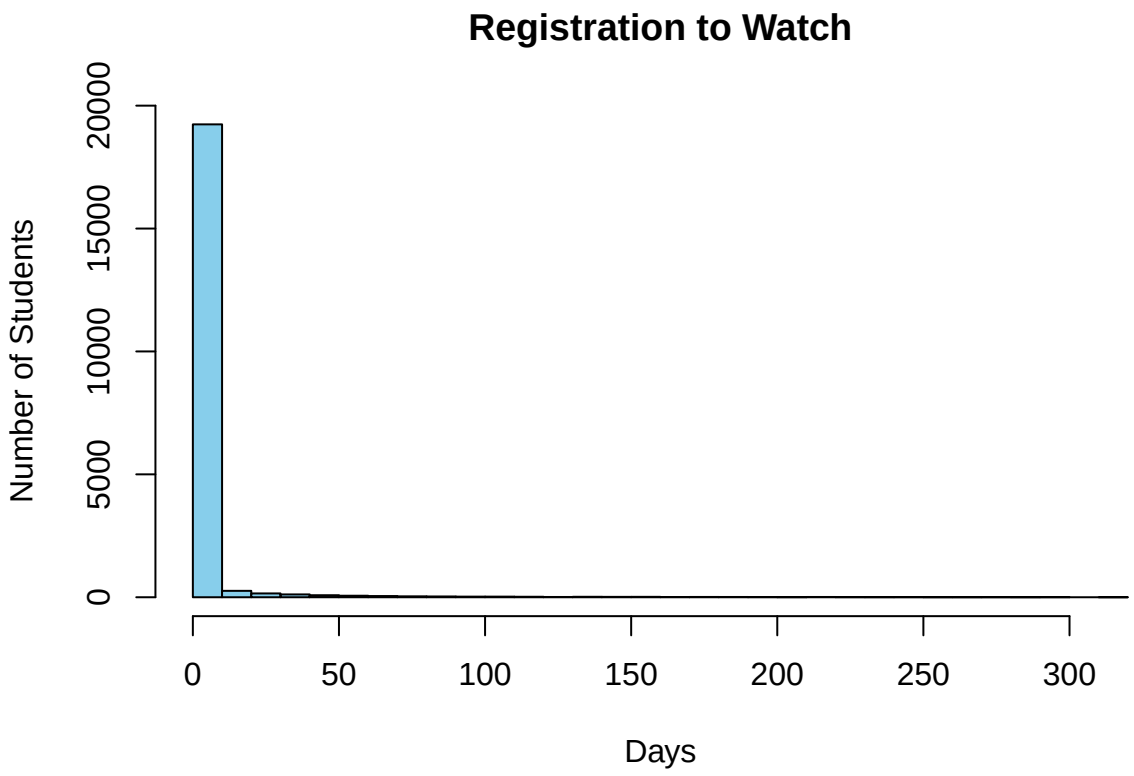
```r
cat("Mode:   ", mode_watch_purch, "\n")
```

```
## Mode:    0
```

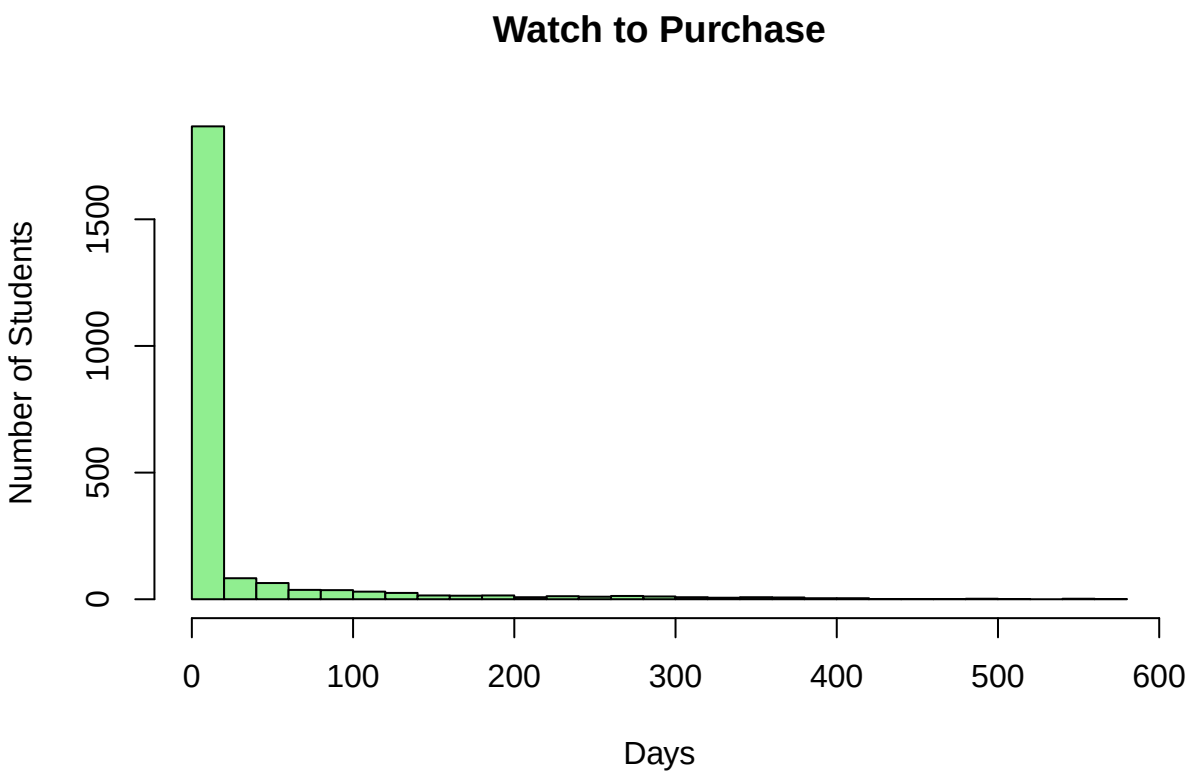
## Interpretation of Distribution Metrics

- **Registration → Watch**
  The mean (3.42 days), median (0 days), and mode (0 days) are all relatively close, suggesting that most users begin engaging with the platform shortly after registering. This indicates a **swift onboarding process** and likely reflects an intuitive and accessible user experience that encourages immediate engagement.
- **Watch → Purchase**
  Although the **median (1 day)** and **mode (0 days)** imply that a majority of users who convert do so shortly after watching their first lecture, the **mean (26.25 days)** is considerably higher. This discrepancy suggests the presence of **outliers**—users who delayed their subscription—pulling the average upward. These delayed conversions may indicate **hesitancy** or a tendency to wait for **special offers or discounts** before purchasing.

```r
hist(
  df$date_diff_reg_watch,
  breaks = 30,
  col = "skyblue",
  main = "Registration to Watch",
  xlab = "Days",
  ylab = "Number of Students"
)
```




```r
hist(
  df$date_diff_watch_purch[!is.na(df$date_diff_watch_purch)],
  breaks = 30,
  col = "lightgreen",
  main = "Watch to Purchase",
  xlab = "Days",
  ylab = "Number of Students"
)
```




## Distribution Insights

Both distributions show that most students take action shortly after the previous step.
The **Watch → Purchase** distribution has a slightly longer right tail, which may suggest that some students wait before purchasing — possibly due to financial considerations or promotional offers.
However, the difference between the two distributions is not dramatic, indicating a relatively similar behavioral pattern across both transitions.

## Conclusion

Most users who convert do so quickly after watching a lecture, as shown by the low median and mode.
However, the higher mean suggests a smaller group delays their purchase, highlighting an opportunity to target these users with tailored follow-up or offers.