# Grammarly Report: February 2016 Insights

Kathryn Meehan, Ph.D.

September 7, 2020

## 1 Introduction

This report covers app usage in February 2016, during which there were

- 4.82 million pings

- 246,840 active users of which:

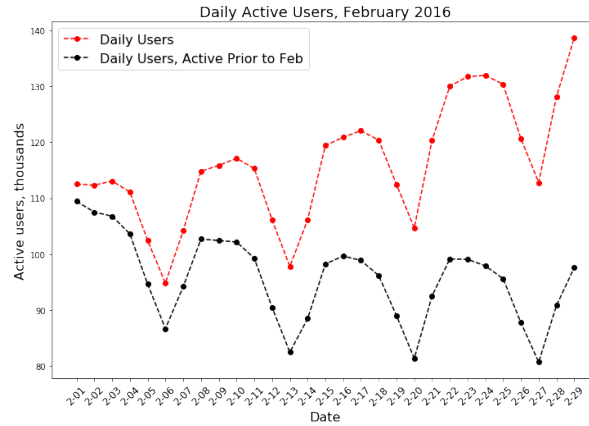  - 86,598 were new (35%)
  - 180,009 were registered (73%)



Figure 1: The total number of active users per day (red) and the number of returning active users- users that had used the app prior to February 2016- per day (black). Note that the y-axis is zero-suppressed.

Figure 1 shows the number of active users for each day of February. There is a clear weekly modulation: app usage dips during the weekend (Friday-Sunday) and peaks during the main part of the work week (Monday-Thursday). Averaging over a seven-day window to reduce sensitivity to this modulation, the week

of February 21-27 sees an 18% increase in the daily use rate compared to the first week of February. Additionally, Figure 1 shows the number of returning active users (users that were active before February 2016) per day. This demonstrates that the growth in daily active users is driven by users that were new to the app in February, and that the retention rate of older users is slightly declining.

## 2 Cohort Analysis by Date

Figure 2 shows the daily retention of users for three different cohorts. Cohorts were chosen by the joining date of the user (the day the user pinged the app for the first time). Three cohorts were formed from users that joined on February 4th, users that joined February 10th, and users that joined February 14th. From Figure 1, remember that February 14th occurs during the dip in the weekly modulation whereas February 10th occurs during the active portion. Despite the variation in the the daily use rate for those different days, Figure 2 shows there is no substantial difference in the fraction of cohort users that are retained once the fraction has leveled off around 40%. This is not surprising since retention should be independent of joining date, unless there is a new version released during the time window of interest or other promotions or changes that would be expected to change user behavior. Promotions to encourage users to join should then be aimed during the higher use periods (weekdays) since 40% of a higher initial fraction will result in more retained users.
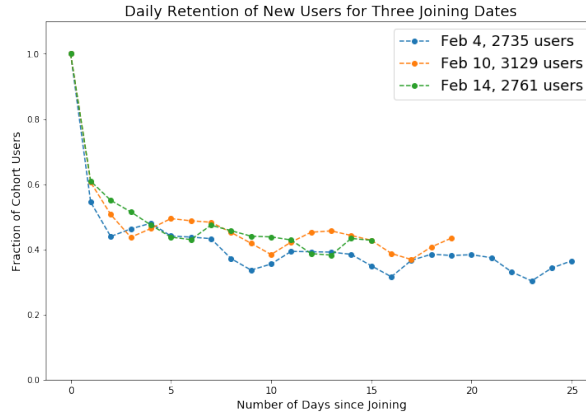


Figure 2: Fraction of cohort users that ping each day. Users are grouped into three cohorts determined by the day of their first ping.

# 3   Cohort Analysis by Source

Figure 3 indicates that the majority of first-time users come from a single source, "Biznesowe+Rewolucje", and that over 90% of first-time users come from just five sources: "Biznesowe+Rewolucje", "answers", "MosaLingua+FR", "shmoop_left" and "program". Focusing on these top five sources, Figure 4 dis-
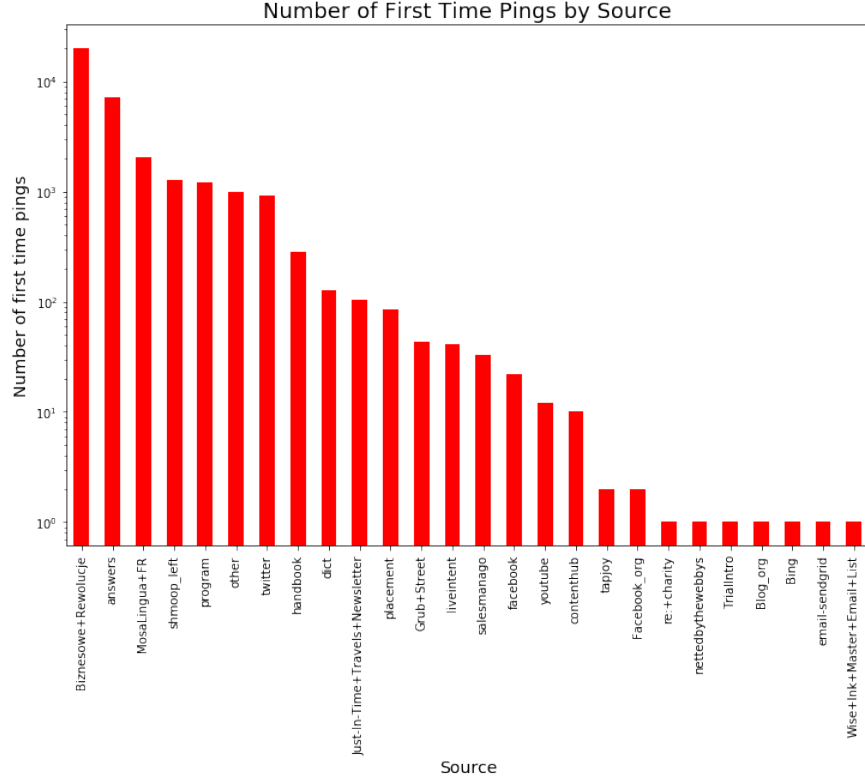


Figure 3: UTM source of the ping. Note that the y-axis is a log scale.

plays the retention curve of February 10th first-time visitors by the source of their initial ping. From this more limited size data set, users from the "MosaLingua+FR" source seem to be the most engaged, although further analysis should be done to determine if this difference is truly statistically significant.

For now, we can turn to a larger data set to help reduce the uncertainties introduced by the smaller cohort size as well as the shorter window time: the users that were active prior to February. Figure 5 shows the top five sources for these users. In Figure 3, "MosaLingua+FR" is the third most common source, yet it is not in the top five sources for returning users. However, since the number of users for sources 2-7 in figure 3 are all approximately the same order of

magnitude, it could simply be a statistical fluctuation that "MosaLingua+FR" appears to bring in more users in February.
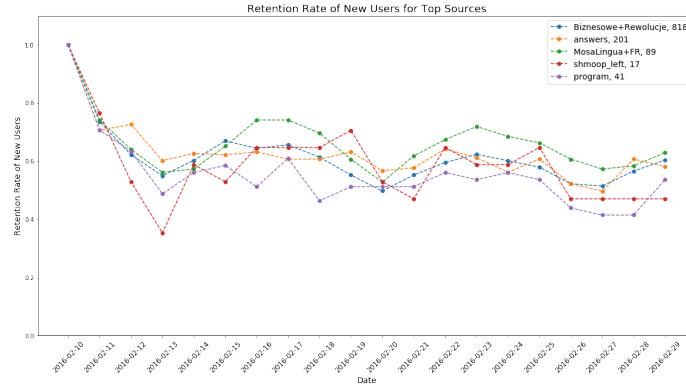


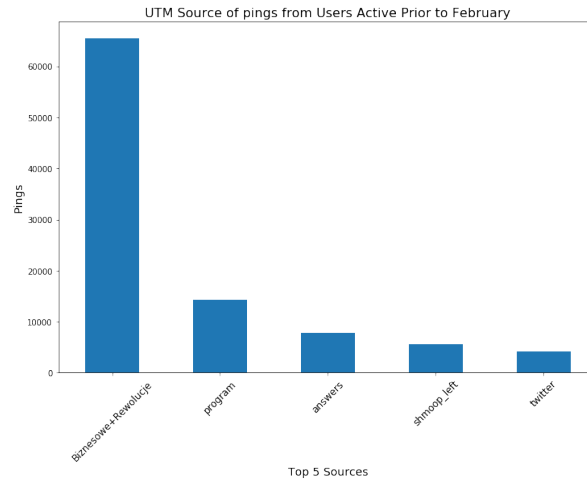Figure 4: Fraction of Feb 10th cohort users that ping each day, by source of their initial ping.



Figure 5: The top five sources of pings for users active prior to February. Note that the y-scale is linear, unlike in Figure 3.

According to Figure 6, the source "shmoop_left" has on average a higher user engagement, by approximately 5 percentage points, than the other top sources. Therefore promotions to attract more "good" users should target "shmoop_left" out of the top five sources.
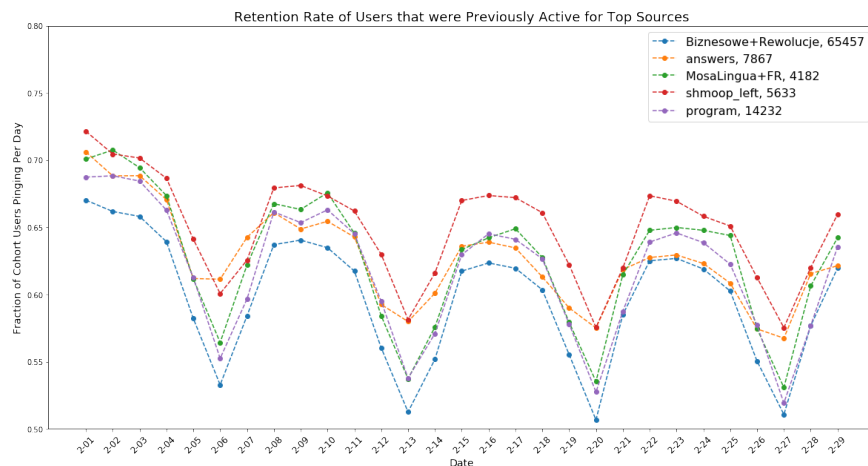
Figure 6: The fraction of returning users that ping each day, by the source of their initial ping. Note that the axis is zero-suppressed to make it easier to see details.

# 4  Conclusions

In conclusion, the increase in daily active users over the month of February is driven largely by new users, so to maintain long-term growth Grammarly should focus on user retention as well as attracting new users. Usage is highest during the peak workweek (Monday-Thursday), so promotions during this period will attract the most users. Long-term retention stabilizes at about 40% of the initial group of new users. However, users from "shmoop_left" average about 60% engagement, and is one of the best sources of engaged users.