App Data Analysis

Introduction

This report outlines the findings from analyzing app usage patterns for the years 2015 and 2016 and includes the 3 app versions.

Purpose of the Analysis

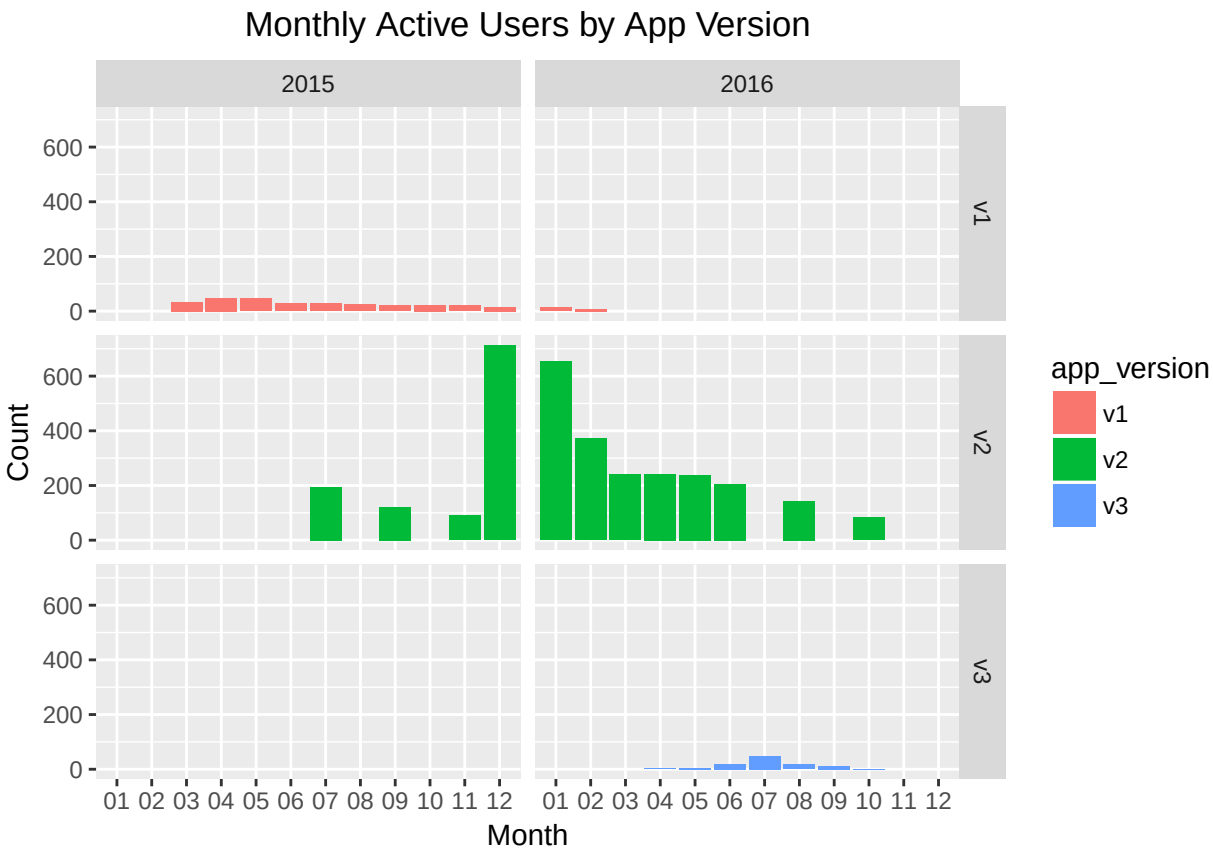
Understanding user behaviour and checking if the high-level KPIs for the application are improving with each release of a new version.

Evaluation of Key Performance Indicators

Monthly Active Users

Calculating Active Monthly Users in this case shows us how many unique individuals used the app more than once in 30 days.The graph below represents that number, summarizing the results for each month (January as 1 and December as 12) for the 2 year span that the analysis was based upon.

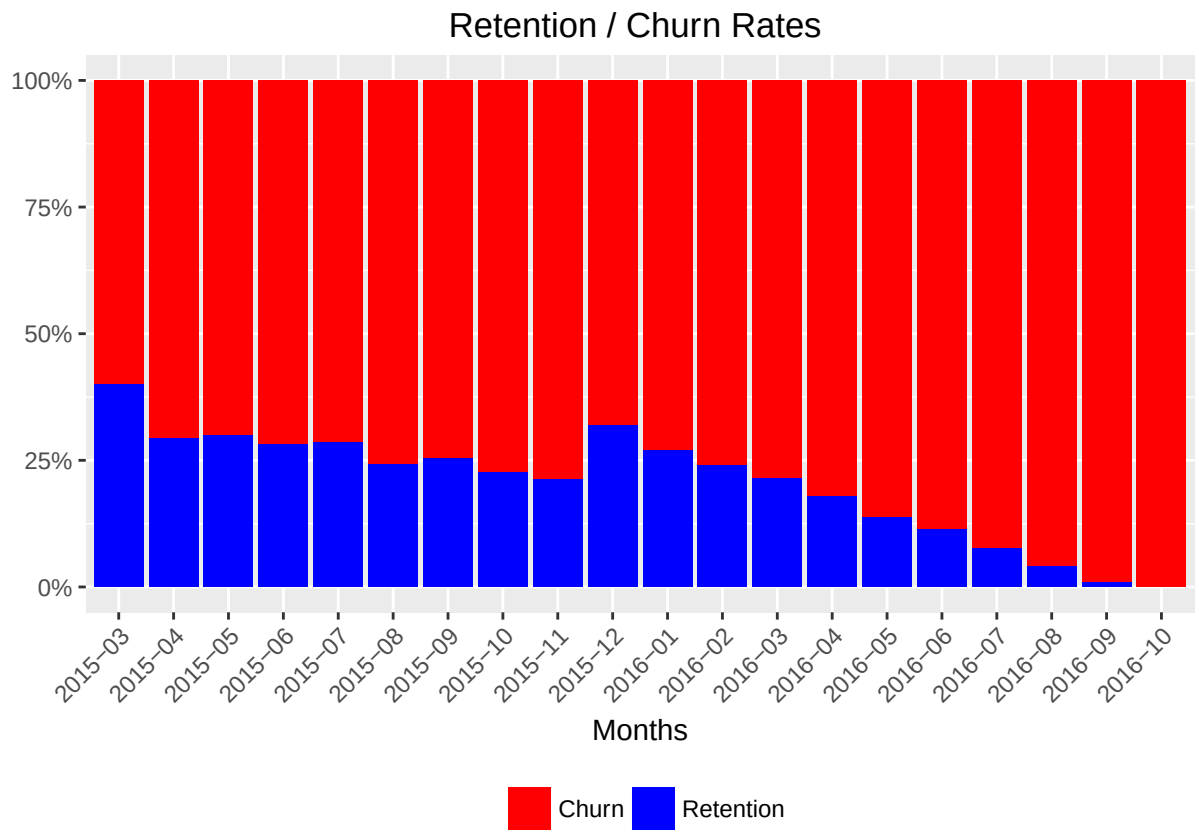
The arrival of the second version of the app promotes better results with increased traffic of new users that tend to use the app frequently. Peak use is happening during the end and beginning of the next year as the transition takes place.



Retention/Churn Rates

A retention rate informs with the percentage of users who are still using the app after installation. The figure can be calculated by counting the number of unique users that trigger at least one session per month if we test for monthly retention rates, then dividing this by the number of total installations.

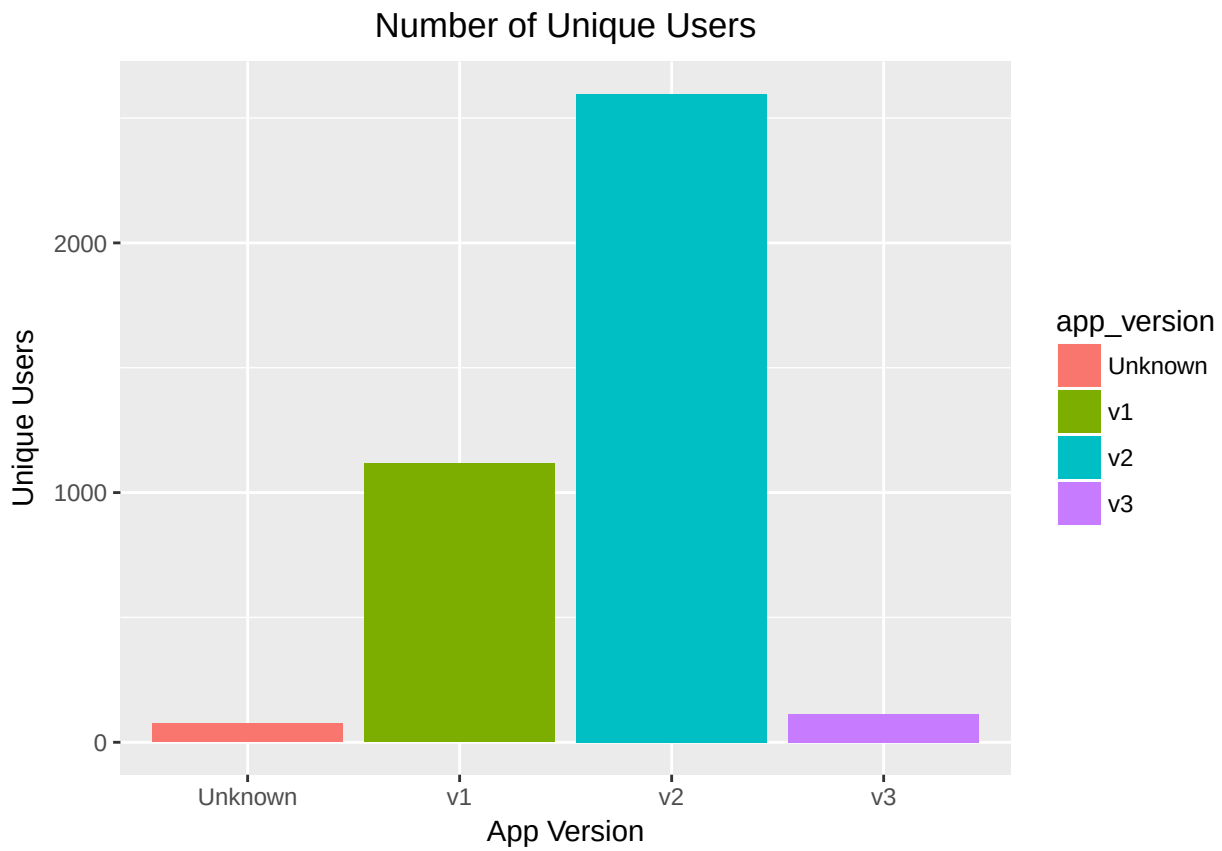
```
## Joining, by = "month"
```



Number of Unique Users

This metric gives an impression of how far each version of the app has reached and provide valuable information as to what can be improved for future updates.

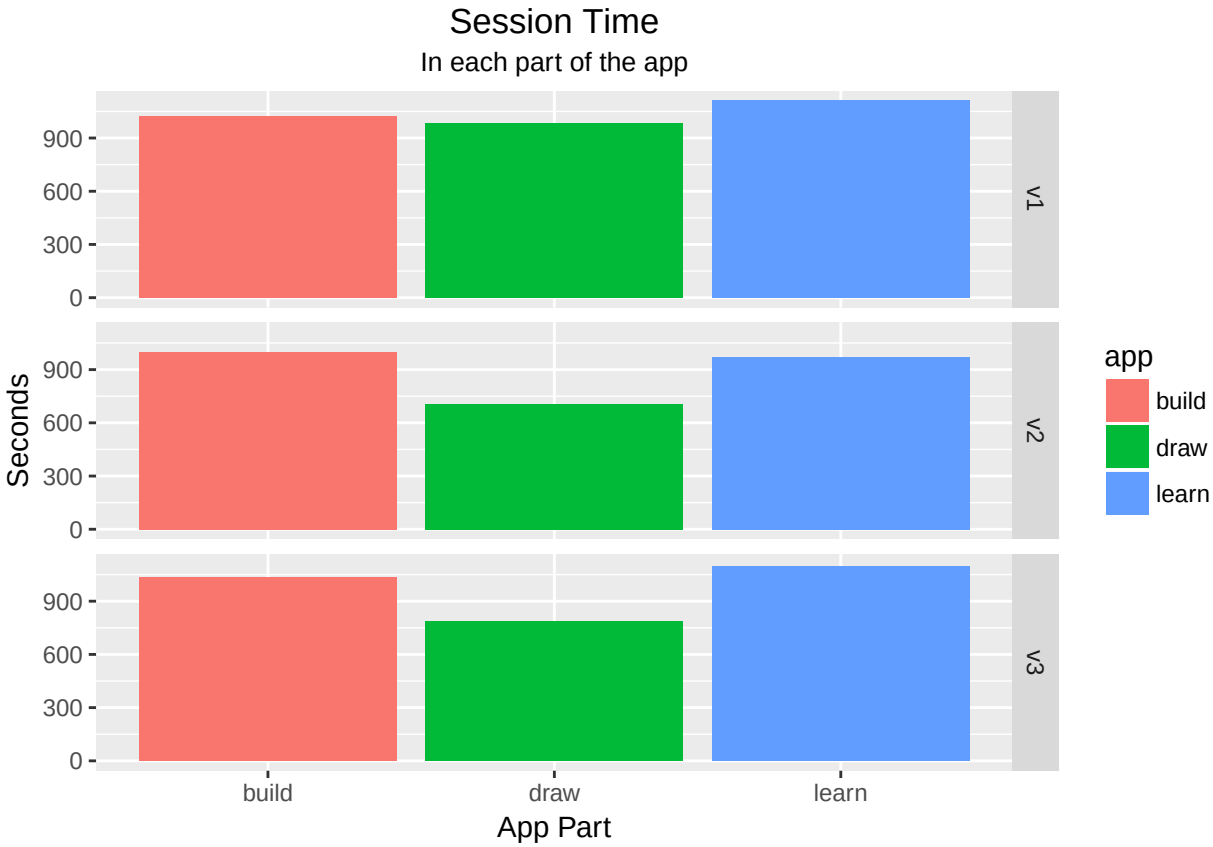
It is evident that the second version did way better than the first version but it is too early for comparisons with the third since the data about this version are only for a 6 month interval while for the second version are more than a year.The first version has more than a year as well. Therefore we can conclude that the second version is successful.



Session Time

At this point a thorough examination of the structure of the app and what is the part that attracts the users is very crucial because its the main purpose that exists.

The following graph displays that in all the three version most time is spent learning and building.



ANOVA for time spent per version

Multiple Hypothesis Testing Between Versions

From the ANOVA test we observed that the p value turned out to be below 5%.Therefore we reject the null hypothesis that states that the means between them are not statistically different in favor of the altrnative.Thus,there is a statistically significant difference between the mean of time of usage per version.

To further discern which mean values are statistically different we performed the Tukey test which does a detailed comparison of every possible combination among the mean session intervals of all the versions.

From the outcome of this test we observe that the mean session times of versions 1 and 2 are statistically different.

```
fit <- aov(w$time_seconds ~ w$app_version, data=w)

summary(fit)
```

```
##                   Df    Sum Sq  Mean Sq F value Pr(>F)
## w$app_version      2 2.087e+07 10434270   6.122 0.0022 **
## Residuals      27723 4.725e+10  1704492
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
TukeyHSD(fit)
```

```
##   Tukey multiple comparisons of means
```

```
##     95% family-wise confidence level
##
## Fit: aov(formula = w$time_seconds ~ w$app_version, data = w)
##
## $`w$app_version`
##              diff        lwr       upr     p adj
## v2-v1 -108.48519 -185.29633 -31.67405 0.0026828
## v3-v1  -52.26581 -177.09974  72.56813 0.5887234
## v3-v2   56.21938  -45.90583 158.34459 0.4006840
```

Methods

Before starting the evaluation, various methods of exploratory analysis occured in order to acknowledge the nature of data,what they mean and how we can ustilize them in the best way possible. Later a check for missing values and outliers detection would complete the screening.The only issues found were some unknown application entries that their effect was minimal.

Results

Mobile apps have become a part of everyday life for smartphone and tablet users. Therefore, there's a large opportunity for businesses to engage their customers and drive more revenue. In the lifecycle of a mobile app, there are 4 important stages: acquisition, engagement, monetisation and retention. Keeping track of these, can improve in the long term the success of the app by focusing on modifying what is observed from past failures of previous versions.

In this report we uderstand that the update from version 1 to the next had a great effect on how many users kept using the app for extended time and that new users reached too. It worth mentioning that the page that deal with learning and building are the interesting to the users and it communicates that the drawing part may need some changes.More data about the third version may appear interesting since the previous showed so much success.