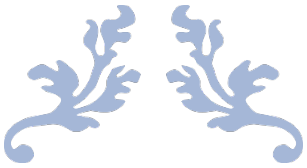


OSXTERN DATA SCIENCE

Xtern Technical Screening

OCTOBER 19, 2017

MANIZ SHRESTHA
Earlham College

Maniz Shrestha, Earlham College
Xtern Technical Screen: Data Science

OS Xtern Data Analytics

The data science team (i.e. Maniz Shrestha) has analyzed the data of OS Xtern usage from **2017-05-01** to **2017-07-31** and have found out many valuable insights for the growth of OS Xtern. This report considers and explains the findings and conclusion of the data analysis.

Basic Information

Data Entry: 25667
Timeframe of the Data: Three months from May to July 2017
Tools used for analysis: Python, Pandas, Mathplotlib and Sci-kit Learn

Analysis

With the basic information and dataset provided, the team made some quick experiments with the data to get a general feel of the data. From this the following information were gathered:

Usage is defined as the number of times OS Xtern has been used by the users. Please not that it is not unique user but rather the users can be repeated if they have used OS Xtern more than once. This is a metric for popularity of OS Xtern.

In this analysis, there are two major trends that the team has been looking towards. The usage of OS Xtern and the number of users that are using it. The following is a glance at the quick facts at a high level:

Number of Users over the timeframe: 100
Maximum usage of OS Xtern in a particular day: 392
Minimum usage of OS Xtern in a particular day: 126
Minimum usage of OS Xtern in a particular day: 279

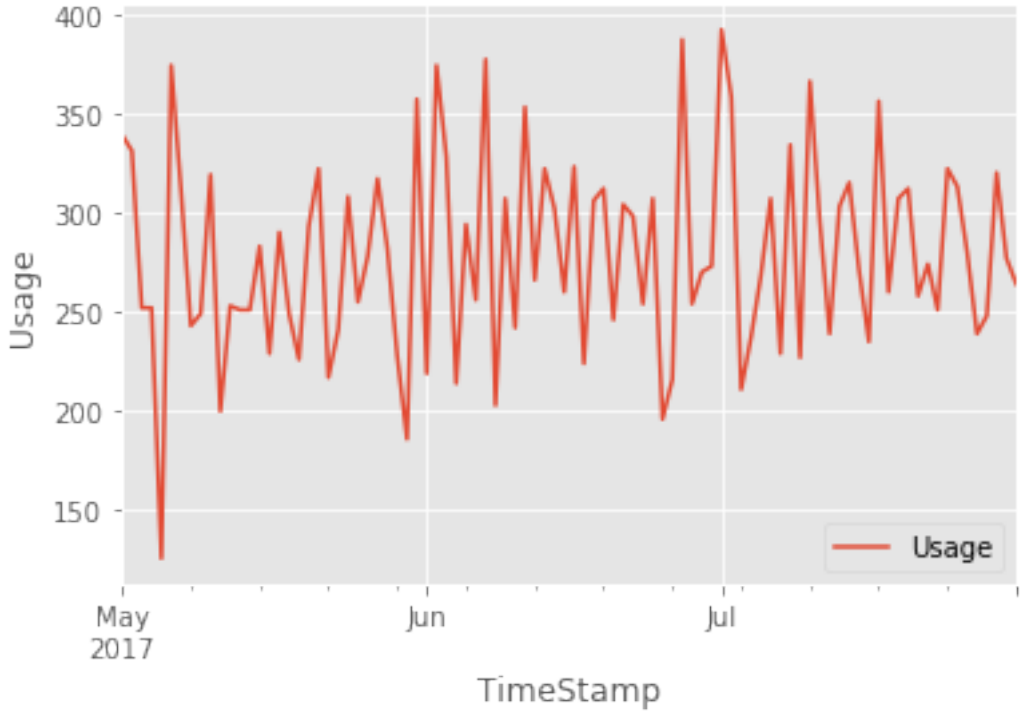
Average Number of Users per day: 98.78 = 99
Average Number of Users per week: 100
Average Number of Users per month: 100

For a detailed analysis, the team has divided the analysis in to three parts: daily, weekly and monthly trend. Please follow the report for more details

Maniz Shrestha, Earlham College
Xtern Technical Screen: Data Science

Daily Usage Trend Analysis

The following graphs shows the usage of OS Xtern on a daily basis from the start of May to end of July 2017.



This figure analysis shows that the usage of OS Xtern is very volatile on a daily basis, something **that we as the developers should be concerned and aware about**. Some questions we should be ask are: **why is there so much jitter in the usage data, are people fluctuation the use of OS Xtern, is there anything we could incorporate in our development to make it less volatile.**

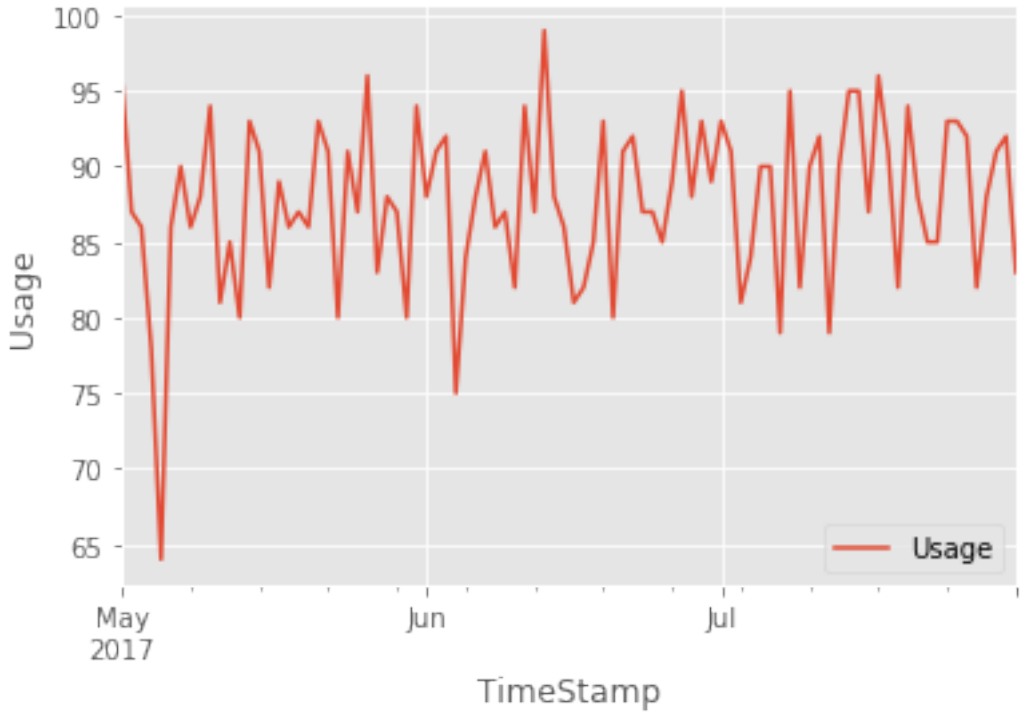
Some more insights:

Statistics	Date	Usage
Peak/Max Daily Usage	2017-07-01	392.0
Minimum Daily Usage	2017-05-05	126.0
Average Daily Usage	2017-05-01 to 2017-07-31	279.0

Maniz Shrestha, Earlham College
Xtern Technical Screen: Data Science

Daily Active Users

Similar to the daily usage of OS Xtern, the data science team has also looked into the number of users using OS Xtern daily to look that the growth in user base of the operating system on a daily basis. The following figure provides us a brief overview:



Similar to the daily usage trend, the daily number of user using OS Xtern is also exposed to **high amount of jitter**. The trend is especially out of place in the first week where it goes to a **minimum of 64 active users** (May be people outside of Xtern were not using it). However, after that period, **the amount of jitter is slightly lower** but the number of users using OS Xtern is still **volatile**. However, one should keep in mind that this volatility is number taking into account the small user base of 100.

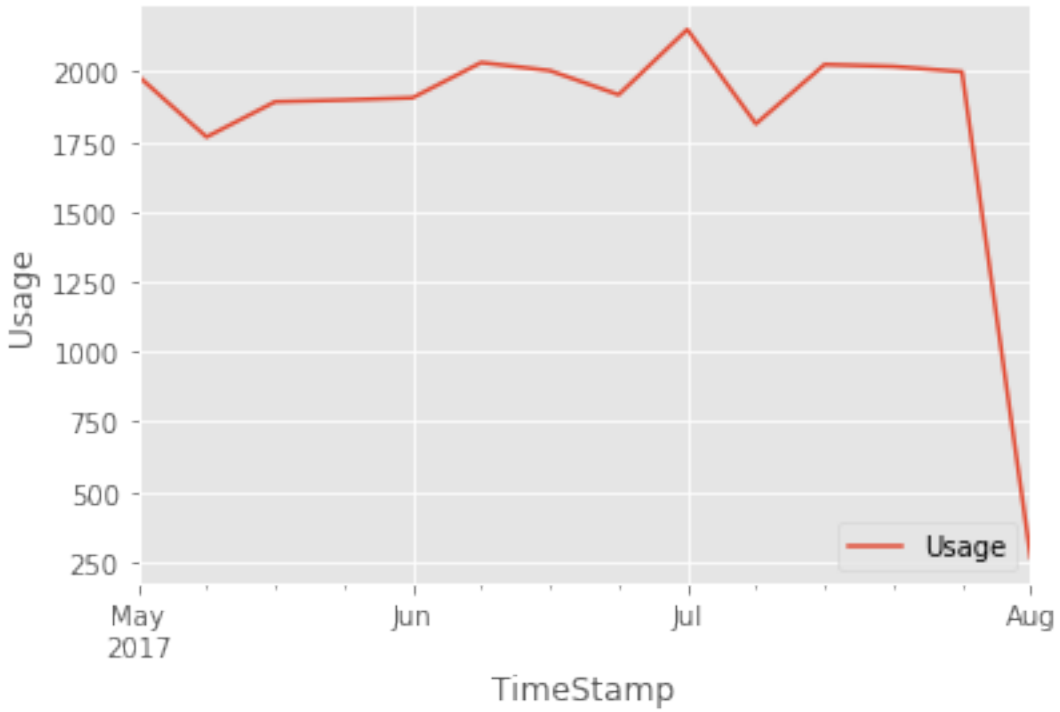
Some more insights:

Statistics	Date	Number of Users
Peak/Max Daily Users	2017-06-13	99
Minimum Daily Users	2017-05-05	64
Average Daily Users	2017-05-01 to 2017-07-31	88

Maniz Shrestha, Earlham College
Xtern Technical Screen: Data Science

Weekly Usage Trend Analysis

The following graphs shows the weekly usage of OS Xtern from the start of May to end of July 2017.



This figure analysis shows that the usage of OS Xtern is fairly consist in a weekly basis except toward the end of July. It is very essential to ask why the weekly usage trend has dropped so significantly. **My analysis for the usage for OS Xtern going down in the last week is that for the last week, there were someday that did not have any data. For instance, the last week contains some days in August but we do not have any entries for August. This must have affected the usage data for the last week.**

It is, therefore, **not necessary** to ask the following questions:

- Is it because our OS is losing popularity?
- Was there a major flaw or glitch found at this time or is there a new competitor to OS Xtern?

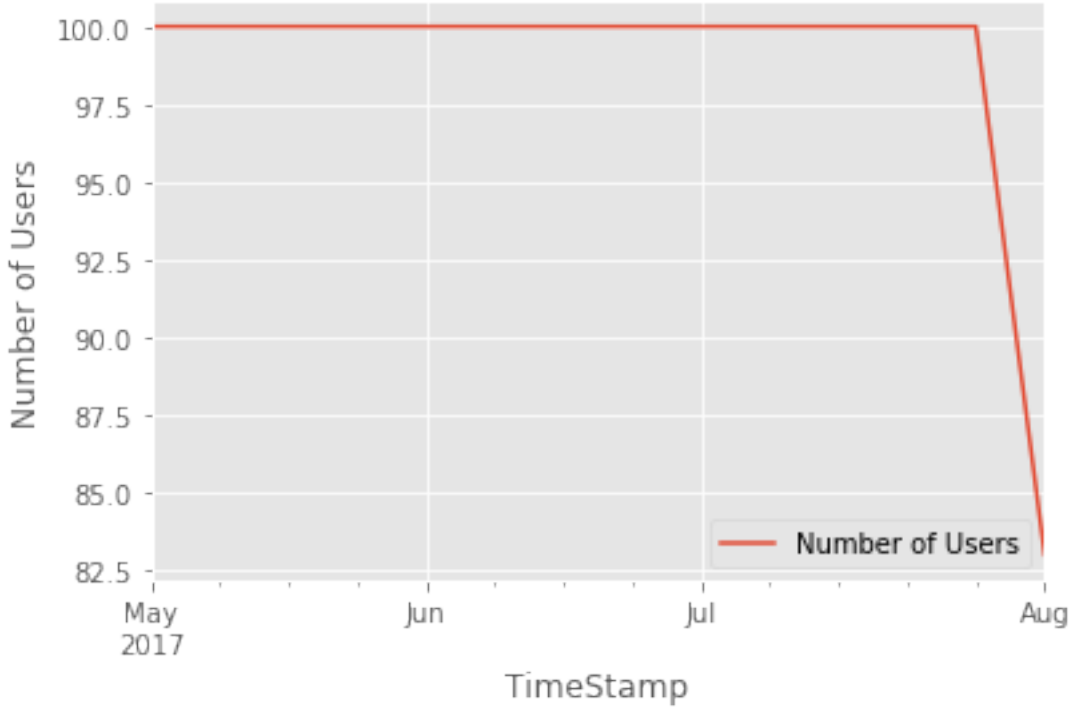
Some more insights:

Statistics	Week	Usage
Peak/Max Weekly Usage	2017-06-26/2017-07-02	2150
Minimum Weekly Usage	2017-07-31/2017-08-06	264
Average Weekly Usage	2017-05-01 to 2017-07-31	1833

Maniz Shrestha, Earlham College
Xtern Technical Screen: Data Science

Weekly Active Users

Note: The decline is not a actual decline as we do not have some entries for August (the last week)



As show in the figure, the number of user on a weekly basis seems to be steady and at maximum for all weeks except the last one. **As with the weekly usage case, we do not have the data for some days in the last week (i.e. the days in August). This should have affected our data and might have left some users who might have used OS Xtern later in the week. Therefore, we can omit the decline in the usage of OS Xtern for the last week.**

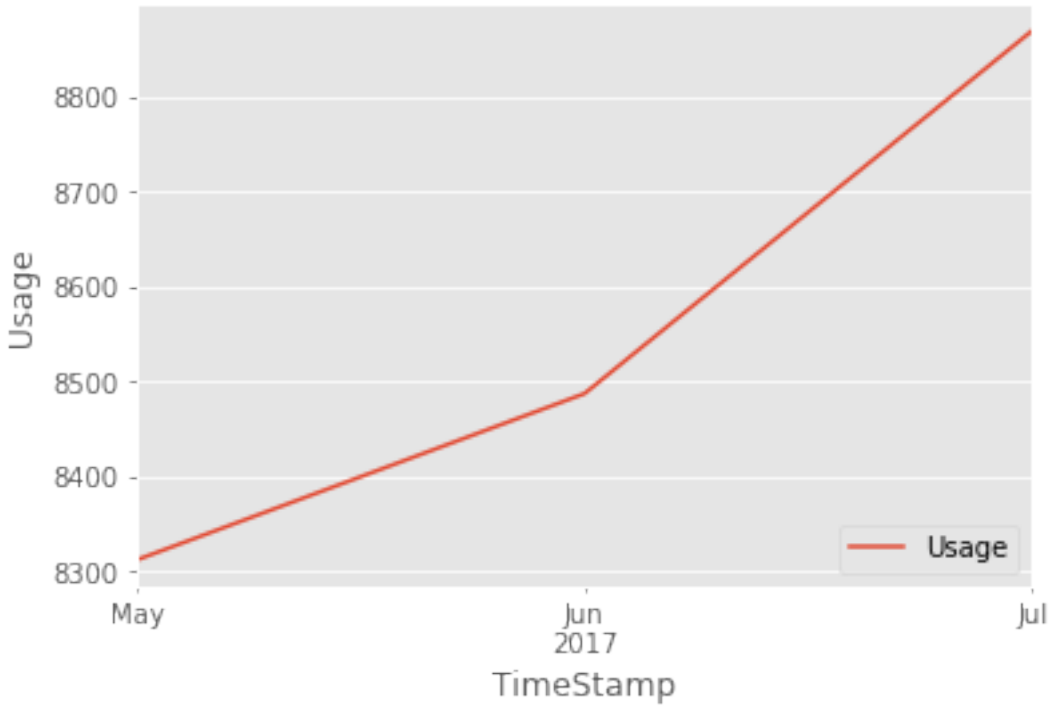
Since all the users are using OS Xtern at least once a week except the last week, **this trend also suggests that there are no users who have been inactive or left using OS Xtern**. This is a major success for Xtern to keep the current user happy and rolling.

However, this is also telling us that the number of user of OS Xtern has not increased since the first week. This means there are no new users. Seems like the marketing folks are not doing a great job.

Maniz Shrestha, Earlham College
Xtern Technical Screen: Data Science

Monthly Usage Trend Analysis

The following graphs shows the monthly usage of OS Xtern from the start of May to end of July 2017.



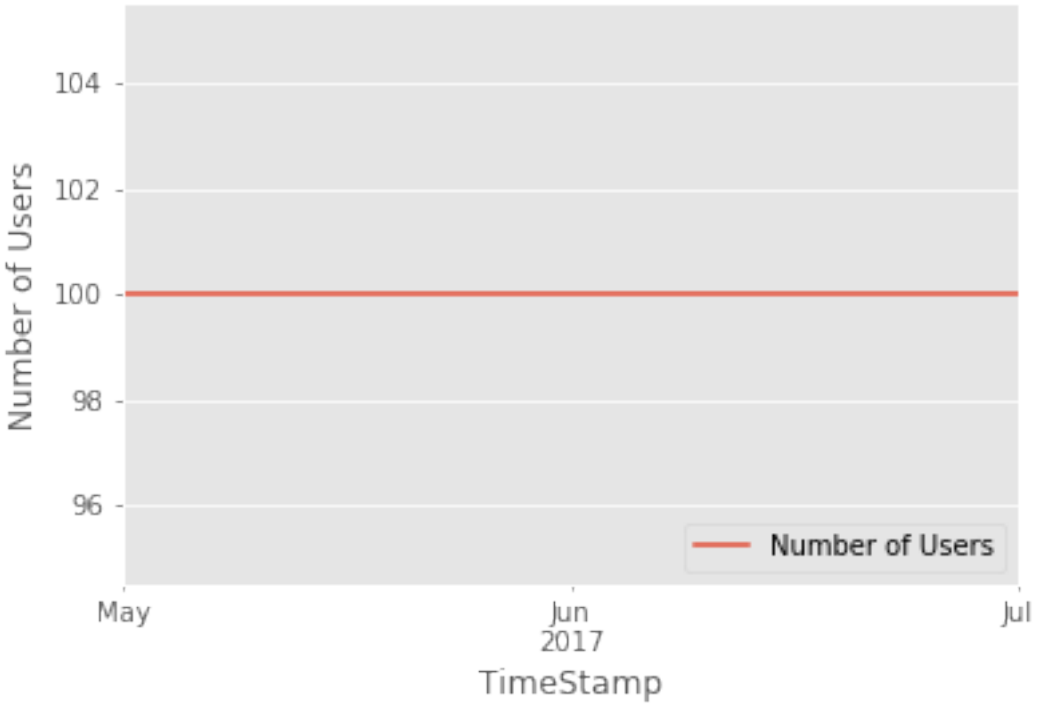
The monthly usage trend analysis looks very optimistic for OS Xtern. Particularly, **it is interesting to note the fact that the rate of growth of OS Xtern is increasing**. So, the usage of OS Xtern is a positive side of the story.

Some more insights:

Statistics	Week	Usage
Peak/Max Monthly Usage	2017-07	8869
Minimum Monthly Usage	2017-05	8312
Average Monthly Usage	2017-05-01 to 2017-07-31	8869

Maniz Shrestha, Earlham College
Xtern Technical Screen: Data Science

Monthly Active Users



The monthly active users is at maximum and consistent throughout the timeframe. **This means that there has been neither growth nor decline in the number of users using OS Xtern.** This was expected as we saw the trend in the weekly user data. This means that the Marketing department at OS Xtern have not been able to attract new people and more effort must be made to spread the word about OS Xtern. (May be they were trying to attract only people in the US Navy)

Bonus: Wow! OS Xtern Users seems to be in US Navy who in the South Atlantic Ocean near Nigeria.

```
locations = data[["xcoordinate", "ycoordinate"]]
fig = gmaps.figure()
fig.add_layer(gmaps.heatmap_layer(locations))
fig
```

Maniz Shrestha, Earlham College
Xtern Technical Screen: Data Science

```
locations = data[["xcoordinate", "ycoordinate"]]
fig = gmaps.figure()
fig.add_layer(gmaps.heatmap_layer(locations))
fig
```

Further Exploration

Due to the limited scope of time and data available, some exploration for a deeper analysis could not be made. Some road maps for further exploration can be using **unsupervised machine learning to cluster data with regards to locations they are used at. Algorithms such as K-Nearest algorithm will be of great use of such exploration.** As shown in the picture, the data points are scattered which means that there might be other pattern within the cluster groups.