Team 3

Dhaval Dholakia

Deepan Sanghavi

Bhakti Chheda

Karan Somaiah Napanda

Rohitpal Singh

pokemon go   
GOTTA CATCH ‘EM ALL ‘DATA’

Case Study 1: Collecting Data from Twitter



Gotta Catch ‘Em All ‘Data’

**Description:** The case study topic that we have adopted for Twitter data analysis is Pokemon GO. Pokemon GO is a location based virtual reality game, developed by Niantic. The game features a story similar to Ash Ketchum, the main face of Pokemon where in each player starts as a novice Pokemon trainer and has to catch pokemons from different locations. The trainer has to walk/move to different locations in order to catch pokemons. It also has other features like Pokemon Gyms where trainers can battle out their pokemons and the strongest Pokemon remains the ring leader and a Pokestop where trainers can collect pokeballs and other items. Another very interesting feature about Pokemon GO is it promotes fitness, such that when a trainer has walked a few miles he is awarded with a Pokemon through the egg he possesses.

Very recently, the game gained tremendous amount of popularity, where in it was being compared to other social networking giants like Facebook and Twitter. A consensus stated that the Pokemon GO app was being used more than any other app, including Facebook, WhatsApp, etc. Pokemon GO has attracted fans from various age groups, from the diehard television show and manga fans from the 90s to the kids from the 21st century.

Pokemon GO is hot in the news not just for its unique game and popularity. It has a name for other negative as well as positive aspects. The game is being used as a marketing tool by a lot of restaurants. The fact that trainers need access to Pokestops in order to restock their items. A lot restaurants are who have a Pokestop in close proximity to their restaurant are using it to attract customers. A pokestop also has another feature which uses a ‘lure module’ on such a pokestop. A ‘lure module’ is said to attract rare pokemons to the respective Pokestop. Restaurant marketers are said to use this to lure pokemons who will in fact lure customers for them. Another example where Pokemon GO was featured in the news for good is a lot of people see it as a fitness cum virtual reality game application. A lot of people have lost a lot of weight due to walking, which includes walking to different places at different heights to find the rarest of Pokemons. But, the game was also blamed for a lot of accidents that was a result of people catching pokemons on highways.

**Motivation:**

Apart from being widely popular, there were several other reasons that proved to be a driving force for the team to decide on to this topic. A major part of the motivation came from the fact that all the team members were huge Pokemon GO fans. We were regular at catching Pokemons and wanted to learn more about the game and the world’s reaction to acquiring a rare pokemon.

It proved to be a worldwide hit as soon as they released the game. A certain survey stated that Pokemon Go’s average daily user will surpass daily users on Twitter soon. It soon became public interest. It had already been downloaded more than a million times. Such public interests were not generated just because of the popularity of the game. It was a result of a few positive and negative views the world had about us. Analysis of such positive and negative feedback added to the motivation.

It’s huge user database also attracted a lot of marketing attention. A lot of organizations were planning to collaborate with the developers for increased profits. An example of such collaboration would be: A pizza restaurant in New York quotes 75% increase in profits after attaching a lure module on the pokestop right outside their outlet.

Pokemon GO was also the reason for a lot of accidents. It is not hard to imagine a person stop in the middle of the road and scream ‘Yeah, it’s a Charizard!’. For example, a man stopped in the middle of a highway to catch a ‘Pikachu’ which led to one of the most serious collisions on highway.

Also, a lot of data regarding Pokemon GO is available on the internet. A lot of people tweet about their captures, their gym battles, and other incidents related to Pokemon GO. Twitter thus provides a huge amount of data which can be used to analyze to produce various results in various domains.

**About the Data:**

The data was gathered from Twitter by using a lot of keywords. The keywords and hashtags particularly used were ‘Pokemon, PokemonGo, Pokemon restaurants, Pokemon fitness, Pokemon servers, Pokemon accidents, Pokemon playing, Pokemon FitBit, Pokemon Restaurant, Pokemon Burger, Pokemon Pizza #DontDrive&Catch, #PokemonGO #playing’.

Tweets were also gathered on the basis of twitter accounts; the follower counts were collected along with the user tweets.

For example, twitter accounts like: Pokemon GO App, Pokemon Fan Pages from various countries (India, US, UK, Singapore & Japan)

A lot of data was acquired which gave results related to the keywords. Some data did not give much meaning to the case study and some did not have the proper format, hence such data needed to be cleaned for further processing and analysis.

NLTK was used to break the tweets into words and to remove stopwords. Regex was used to clean the tweets on the basis of hashtags, RTs (retweets) and punctuation.

Irregularities in data were removed. Regularities such as typing error or irrelevance were discarded.

Following are the details of the extract:

1. Number of tweets: 12,000.
2. Number of followers: 116,000 and counting….
3. Size of the tweets: 33.5 mb of data was generated

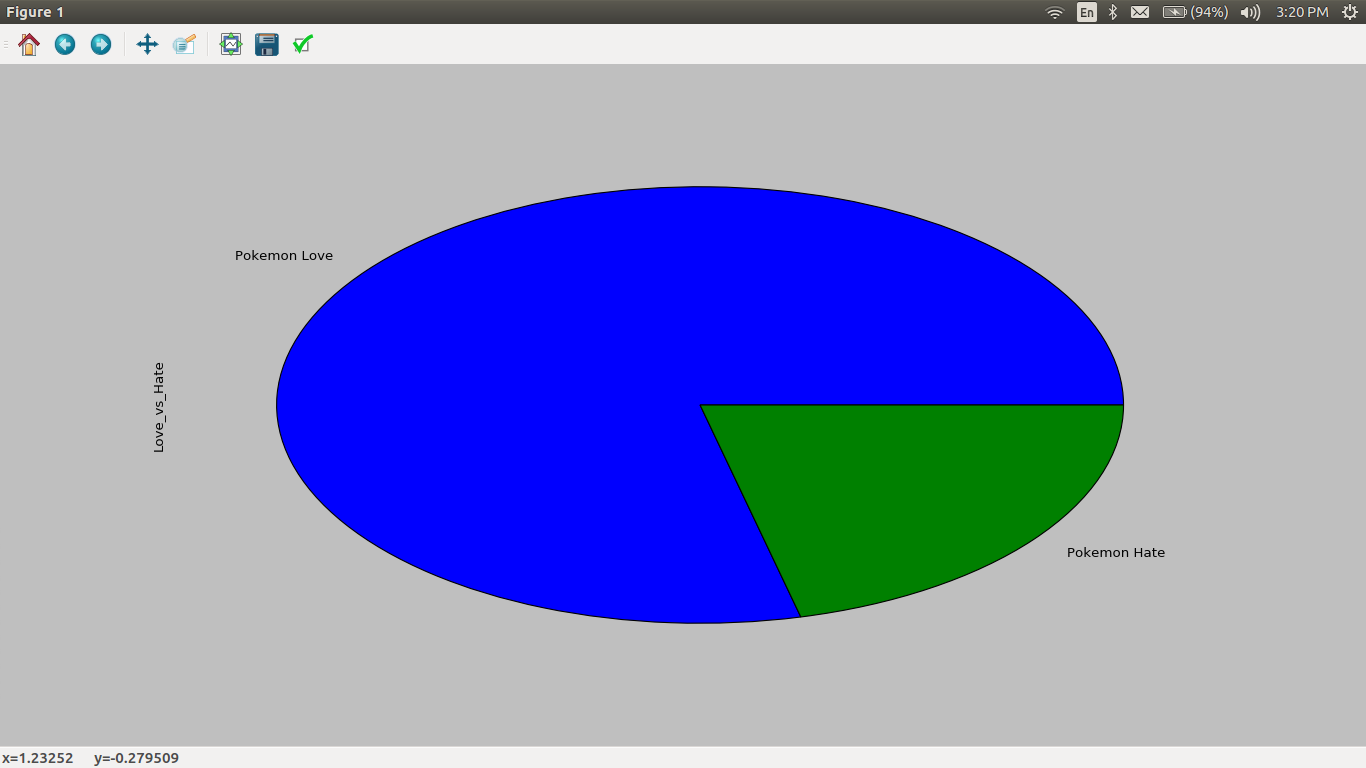
The data that has been gathered is on the basis of the particular user, the number of tweets, the followers, the location, time zone, favorites and retweets of different users.

**Analysis on the Data:**

After cleaning the irregularities in the data, the data was analyzed. The analysis was carried out using Python and its corresponding libraries. For visualization, Pandas and MatPlotlib were used. Moreover, ‘nltk’ and ’regex’ were used for processing the data.

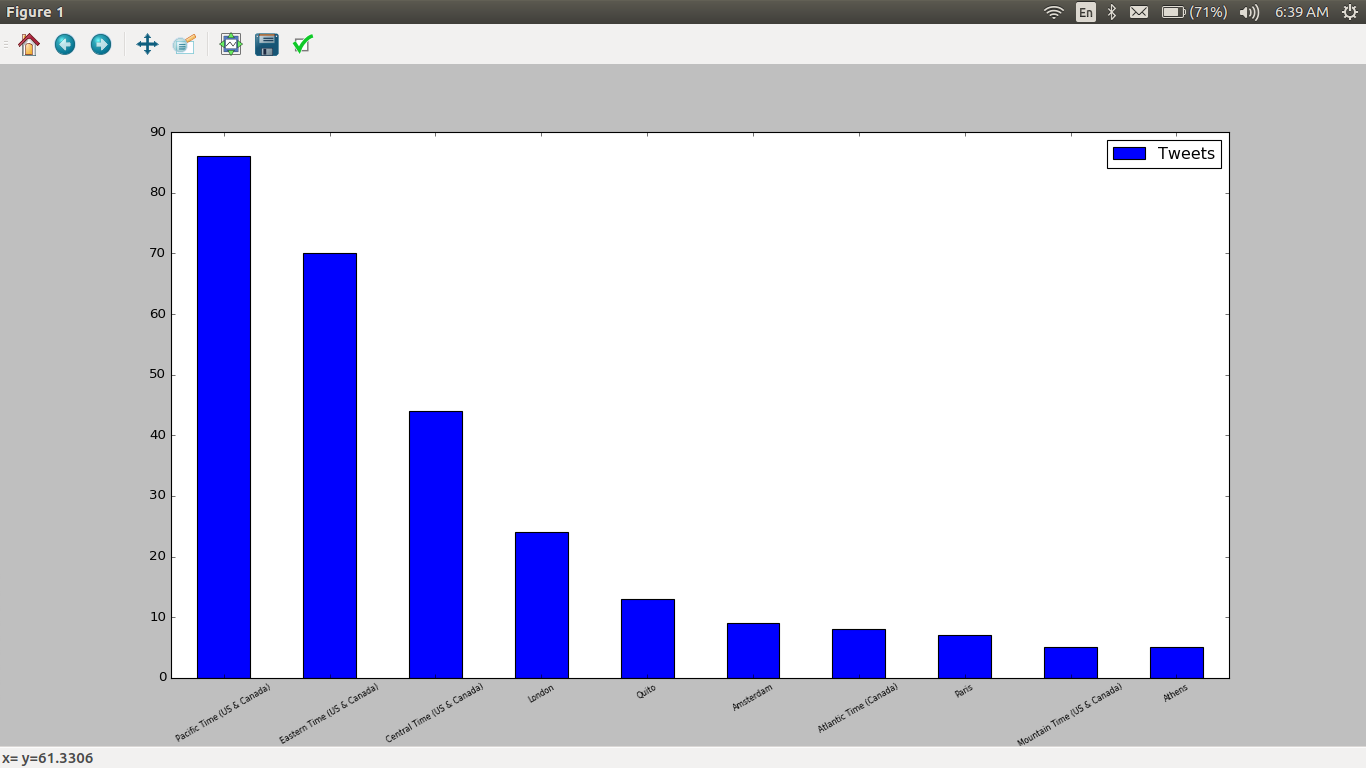
The different kinds of analysis produced are as follows:

1. Analyzing the popularity and fan following of Pokemon on Twitter. As we all know, Pokemon GO has a lot of followers and huge fan base. We tried to gather data from Twitter that would help us estimate the popularity of Twitter.



The above pie chart shows a separation between tweeting for and against Pokemon GO on Twitter. Clearly, it indicates that a majority of the tweeters are for Pokemon GO

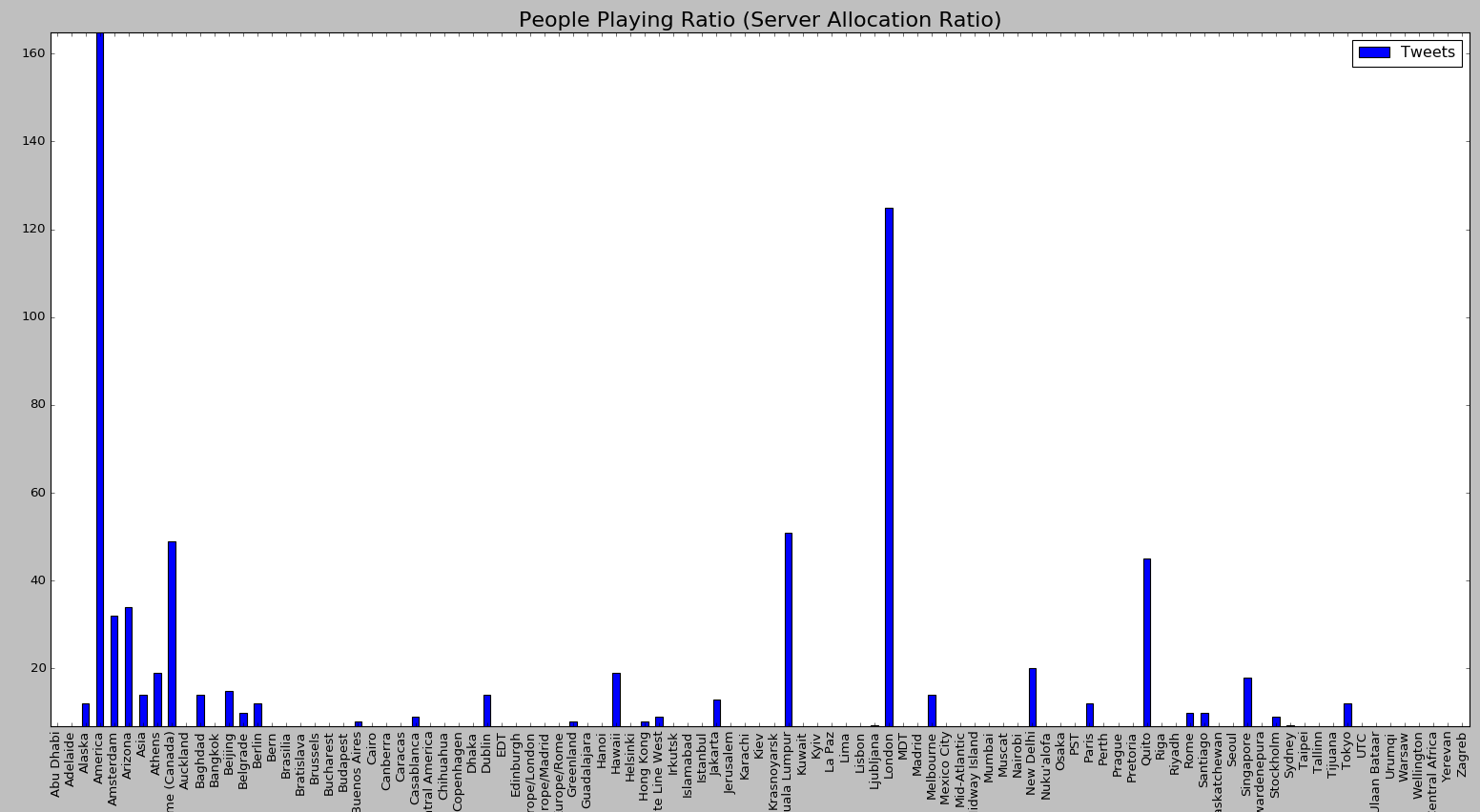
1. Data acquired through extraction of major keywords like ‘Pokemon Accidents’, ‘Pokemon Rob’, ‘Pokemon Cliff Fall’, ‘Pokemon Injuries’ and also we used hashtags like #DoNotCatchAndDrive. Queries for this produced results (tweets) that were related directly or indirectly to accidents. These tweets either reported accidents or any incidents that were occurred due to people playing Pokemon GO. A lot of tweets were news channels reporting the incident. We also gathered few tweets by filtering the news accounts and PokemonGo keywords.



The above graph depicts a general trend by location where the number of tweets that are related to Pokemon GO and accidents. The United Stated tops the chart with the number of accidents cause due to distraction by Pokemon GO. Such data can be used by Niantic to improve their pokemon spawning mechanism and pokestop locations to avoid such incidents. Thus, in the new update, Niantic ensures if the trainer is driving or not and it has removed pokestops and pokemons which are nearby highways.

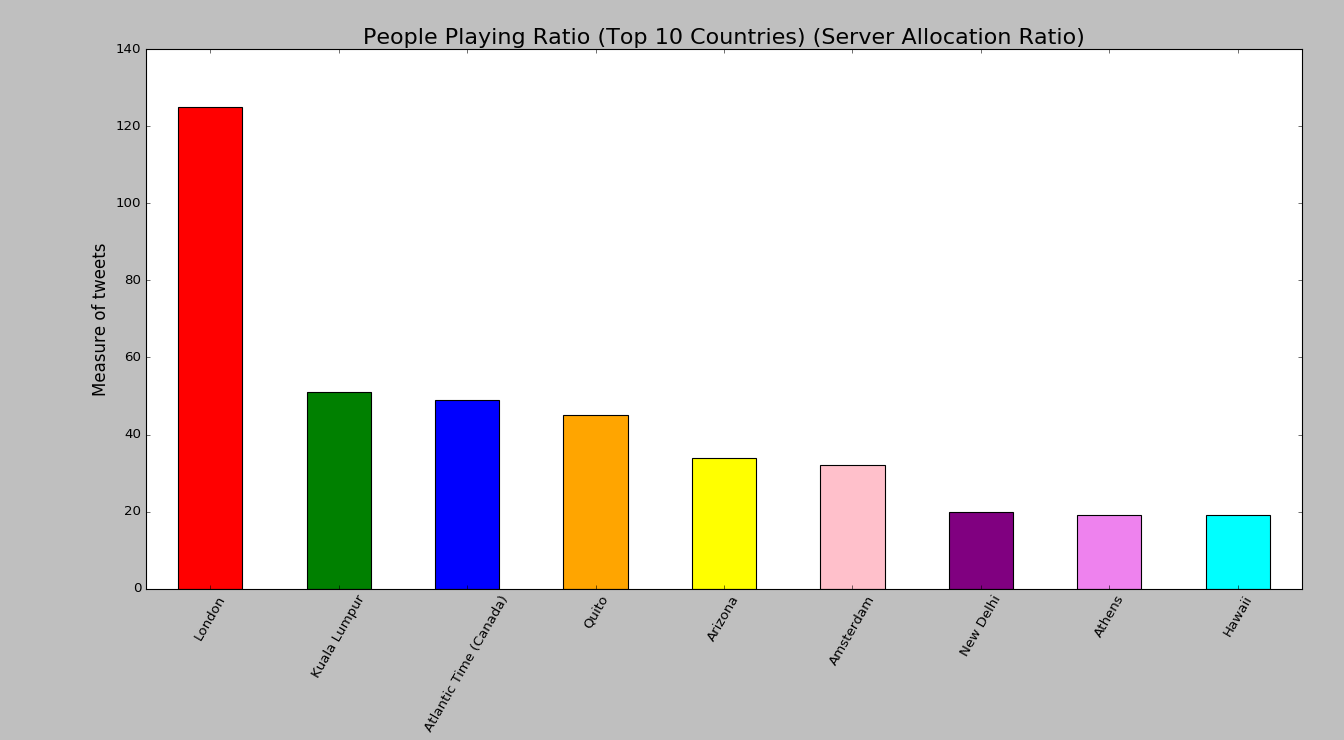
1. A problem faced by Pokemon Go developers, Niantic was due to registration by a lot of users around the world. They did not have enough servers to handle the load. Server failure was faced all over the world by users from different continents due to increasing number of playing users. We also analyzed server issue based on tweets and issues.

The following graphs show where the server load is more, which countries are top competitors in the number of people who are playing the game.



The above graph shows that America leads in the number of Pokemon GO users and hence is the most susceptible to server failure. After America, the United Kingdom comes second followed by Kuala Lampur and Canada. Thus, it illustrates that the American continent is most susceptible to server failure.

Figure below shows the general trend of the number of tweets where people tweeted more about Pokemon GO. It’s the summary of the above graph which shows the most critical places where there can be server issues



1. Analysis of the release of Pokemon GO Plus bands was also done. Analysis was on the basis of comparison with other fitness bands like ‘FitBit’. We found that a lot of people compared FitBit to the Pokemon GO Plus band and a few of them even suggested a collaboration between the two which would have increased FitBit sales.

Around 3,469 tweets were collected that were linked to the hashtag ‘#PokemonGOPlus’. This data was collected in a week’s time.

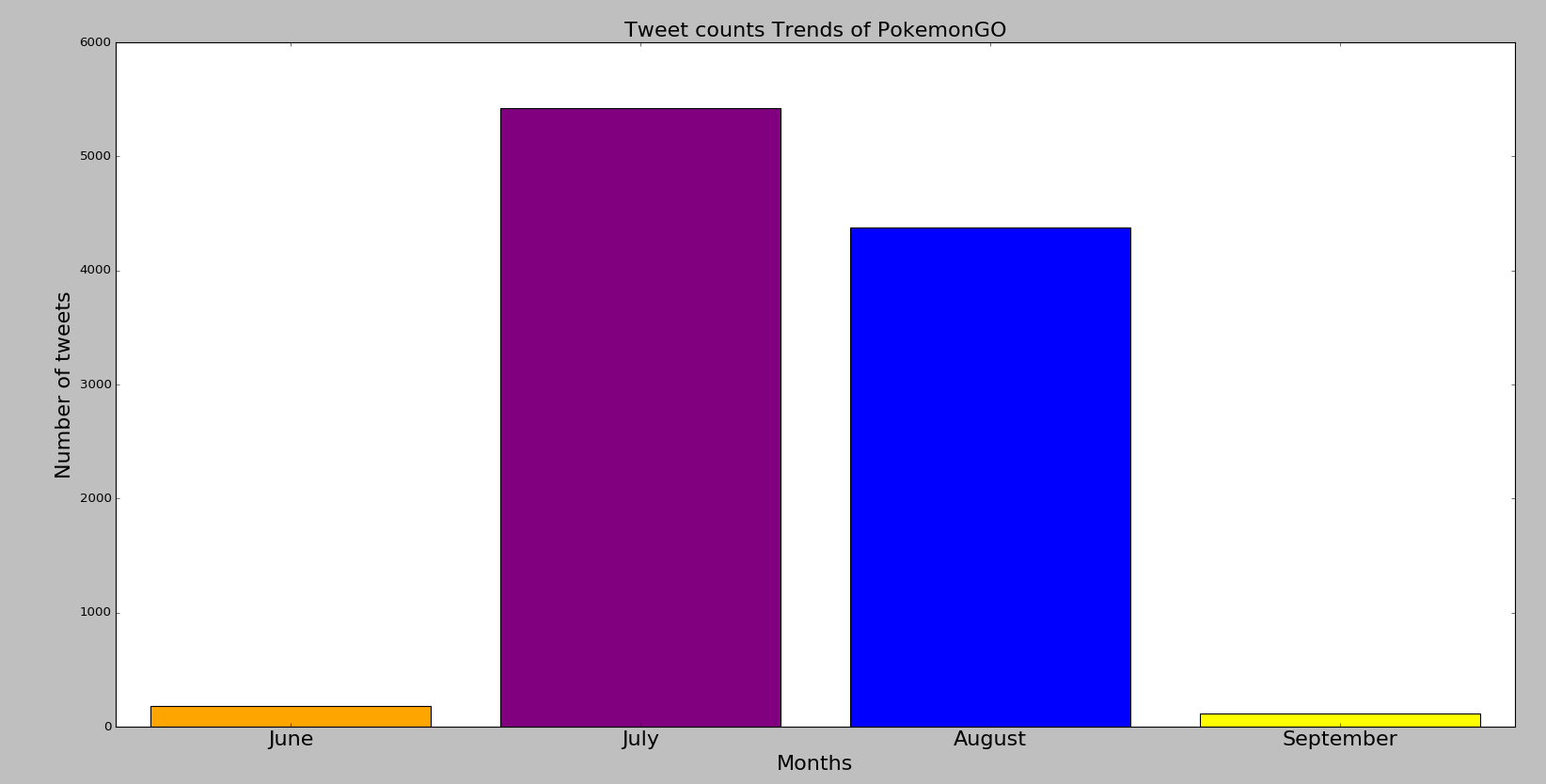
A lot of restaurants have also collaborated with Pokemon in some way or the other. The collaboration is such that ‘lure modules’ are planted by restaurants on Pokestops to attract customers. A lot of restaurants reported a profit of upto 75% in their sales.

We have come across a lot of situations where restaurants have adapted to marketing strategies using PokemonGO where in restaurants have used lure modules to attract pokemons to the pokestop which will lure customers for them.

Another example is where a lot of restaurants have adapted to making products that a pokemon look a likes. A café in Australia has started to make Pokemon shaped burgers. A California based outlet has named a shake after a Pokemon.

**Findings of the Data:**

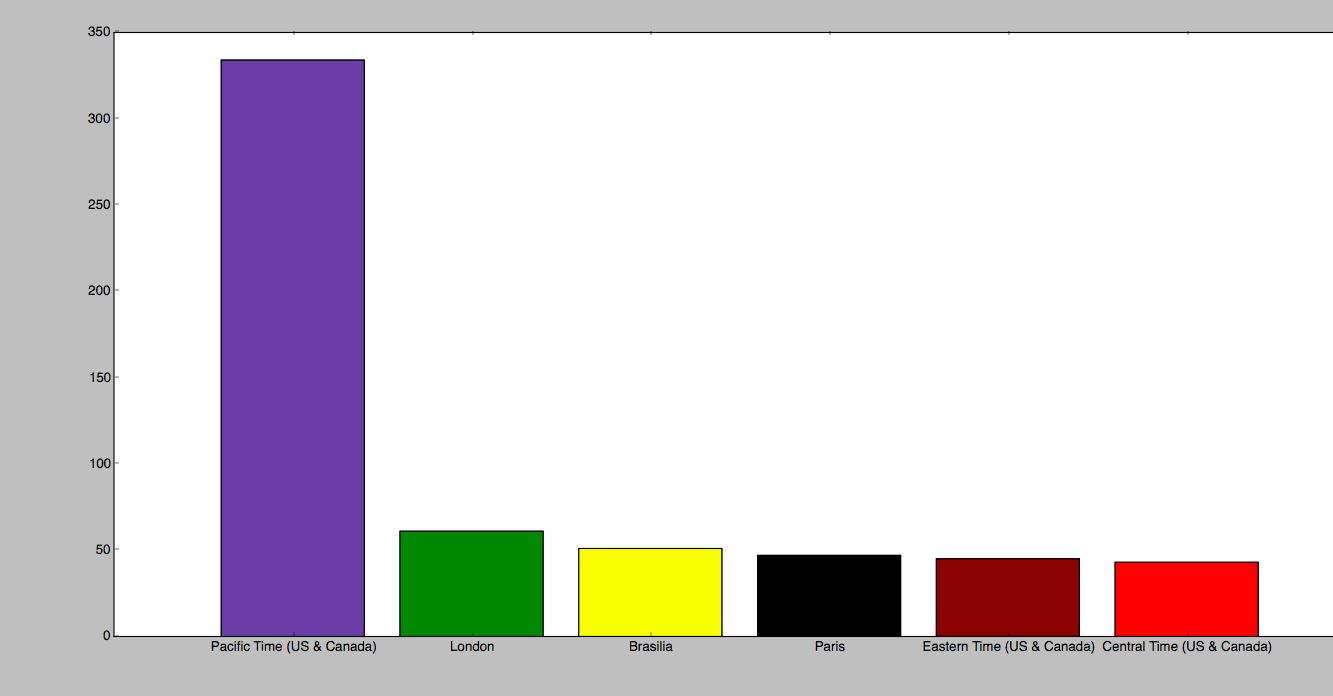
After analysis of these tweets, the most interesting observation we have come across is that the interest in Pokemon GO was sudden and it feel to a great extent. The number of tweets associated show a steep decrease in the interest of the users. Before the release of Pokemon GO, in the month of June there was average excitement of the game. We could observe a sudden increase in the number of tweets in the month of July, when the game launched. The number of tweets again reduces a little in the month of August to face a sudden decrease in the month of September. Thus, we can conclude that the public interest is decreasing and that Niantic needs to come up with a strategy in order to keep the interest going.



Parameters that can benefit the Pokemongo App:

i. Servers Location :

Based on data , heavy users are from Pacific Time (USA & Canada) timezone in proportion to other locations.

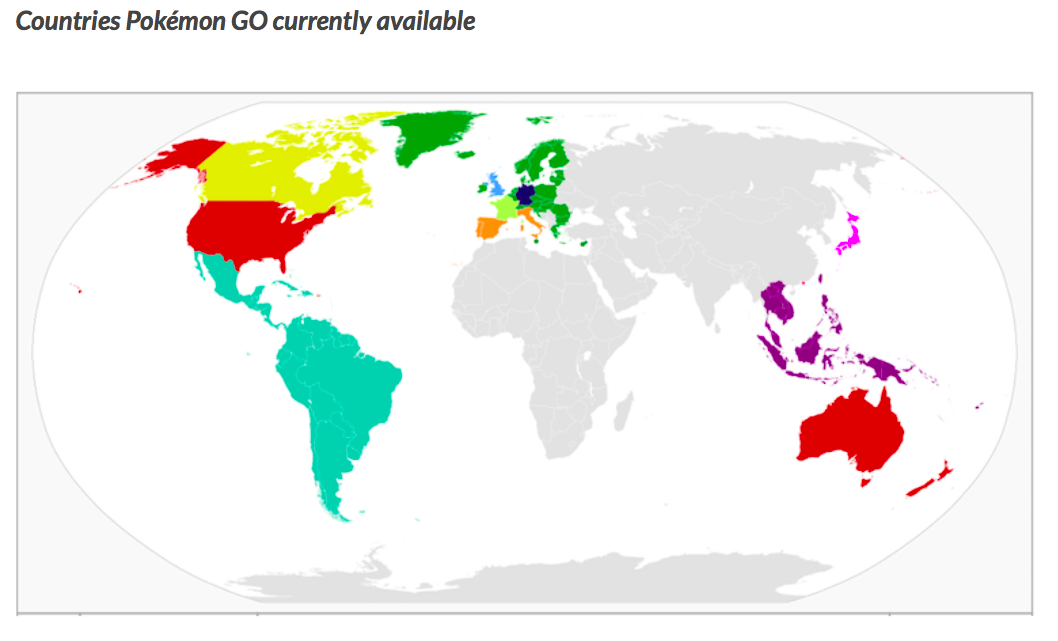


A surge Increase in the usage of Pokemongo App at particular location break down the server. So distributed server location at these locations is must .

With increase in the distributed server location, more users can access the app simultaneously .

Result : Company must invest in servers locations.

ii.



Nintendo, a pokemon based company is officially available at major parts of the world such as USA, Australia, Uk etc. But PoekmonGo is still to be released in major parts of Africa and Asia.

Countries that have diverse markets:

Considering Asia, majorly India and china is having diverse markets in terms of other social media apps such as Facebook, Quora etc

Without officially availability of PokemongoApp in India, many users are accessing this game and surge is exponentially high.

Result : Investing in diverse market will definitely help the Nintendo to increase the sales.