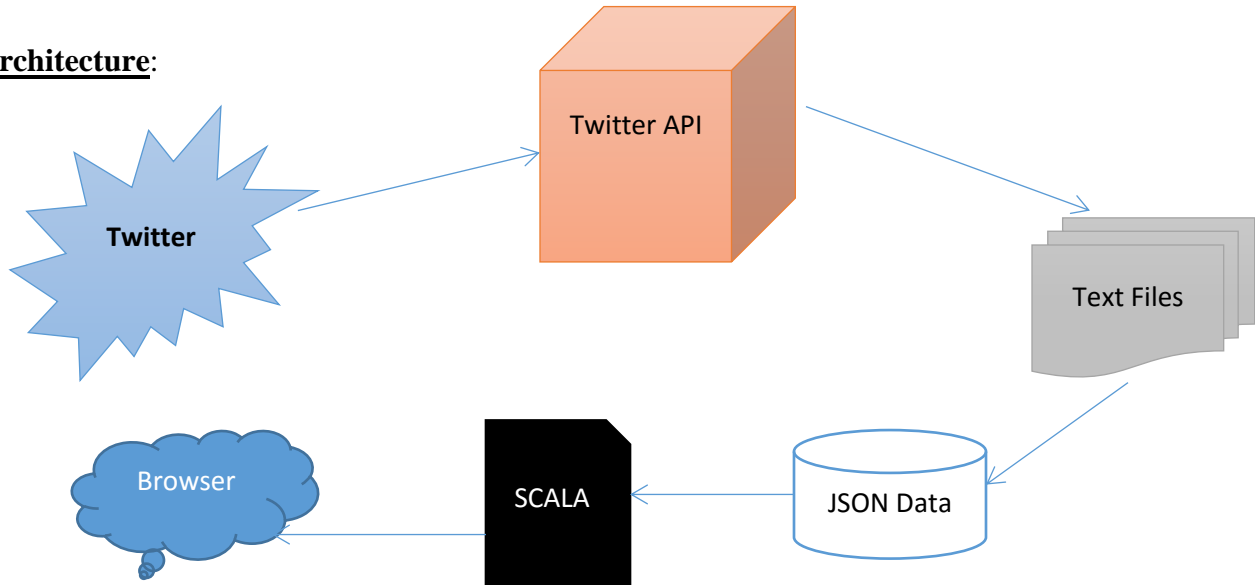


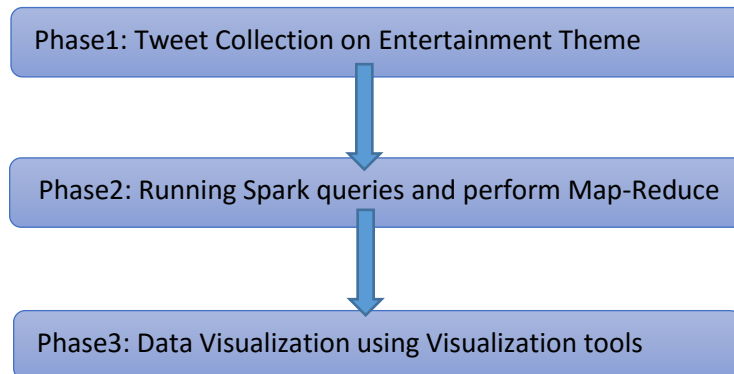
## Twitter Project

### Phase 2

#### Architecture:



#### Design/Phases:



#### Data Collection on Entertainment:

- Sports
- Movies
- Music
- TV
- Adventures

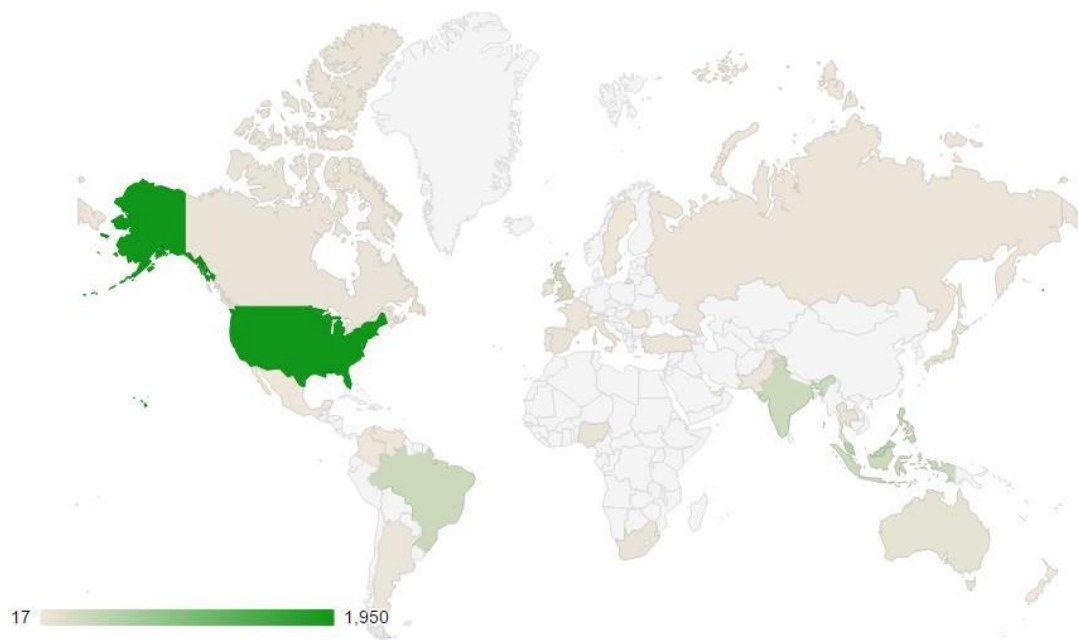
### Tools and Technology:

- Tool: Spark on IntelliJ(IDE),
- Technology: Scala, JavaScript, HTML

### Queries:

*Country with more number of tweets:*

SELECT place.country, count(\*) as country\_count FROM tweets2 where place.country is not null GROUP BY place.country Order By country\_count desc limit 30



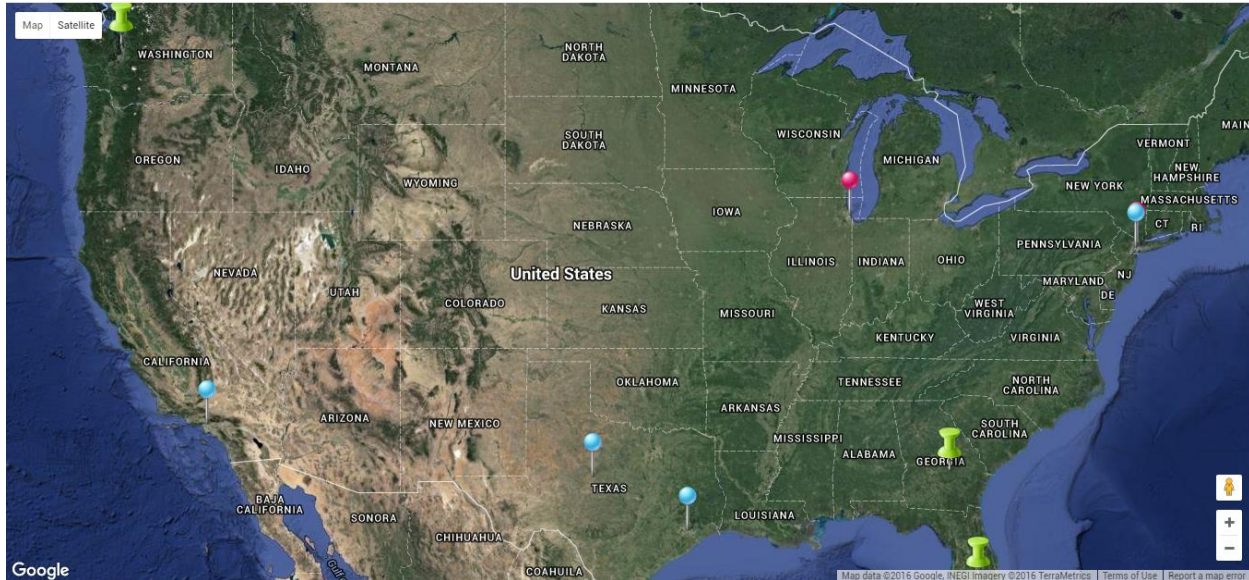
*Top 10 retweeted tweets:*

SELECT text, count(retweet\_count) as c FROM EntertainmentTable group by text order by c desc limit 10

```
[RT @CarAutoDaily: twenty 1 pilots: Dwelling of Gold [Official Movie]... - https://t.co/xlzCSv1Tdil
[RT @SportsCenter: Kobe is on 8Y"¥8Y"¥8Y"¥8Y"¥! He's got a game-high 22 Pts on 10-13 shooting. https://t.co/Qm2RL1vr9E1
[RT @CheapVacationss: 11 Mind-Blowing Facts You Probably Don't Know About Disney Movies https://t.co/QC6tiFWqOV https://t.co/5MD0i9TKDb1
[RT @ThatsSarcasm: i wish i had that friend that lived next door and they'd come over just to kick it and watch movies or something]
[RT @FindAllMovies: Childhood Fantasy Destroyed By The Sinister Origins Of These 15 Disney Movies! https://t.co/Jr3u8vnlwq https://t.co/MTOzâ€¦]
[RT @NBA: #ThisIsWhyWePlay https://t.co/Se939Bw7J31
[RT @FindAllMovies: Access the largest online streaming database, Instantly watch your favorite movies online https://t.co/EUW2c5QG7q https://t.co/â€¦]
[RT @FindAllMovies: Watch movies online for free, stream movies fast and easy and download full movies https://t.co/EUW2c5QG7q https://t.co/â€¦]
[RT @FindAllMovies: I Speak Fluently in Movie Quotes and Sarcasm 15 Oz Stainless Steel Travel Mug https://t.co/e3hzQ7F5XX https://t.co/CHPulâ€¦]
[RT @FindAllMovies: Watch TV series and full length movies on demand here https://t.co/Dz5CAgSsvI https://t.co/7VjW56Kd0g]
```

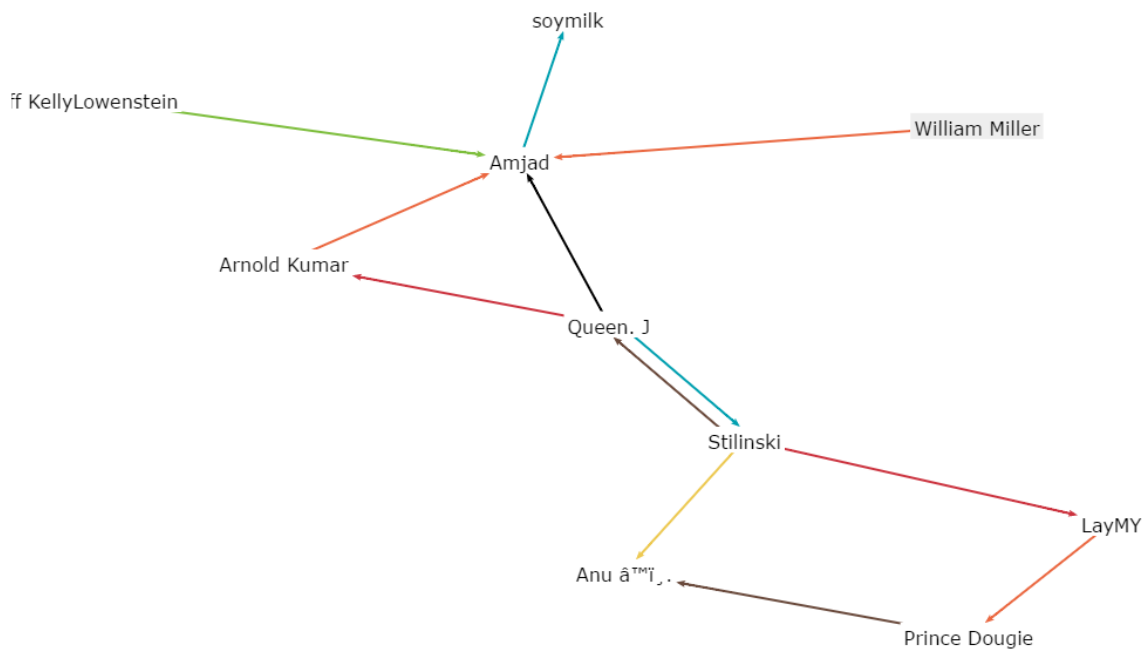
***Top places tweeted more about movies in United States***

```
SELECT place.full_name, count(*) as c FROM EntertainmentTable
WHERE text like '%movie%' and place.country = 'United States' GROUP BY place.full_name
ORDER BY c desc limit 15
```



***Top 10 most followed users***

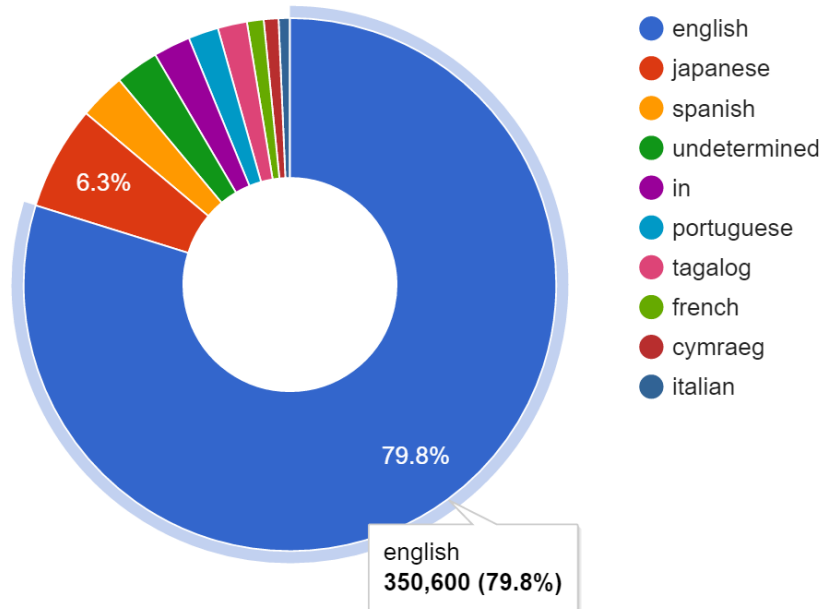
```
SELECT user.name, user.followers_count FROM tweets2 ORDER BY user.followers_count
desc limit 20
```



***High preferable languages used to tweet***

```
SELECT lang, count(*) as c FROM EntertainmentTable WHERE lang is not null GROUP BY  
lang ORDER BY c desc limit 10
```

In which language did you tweet?

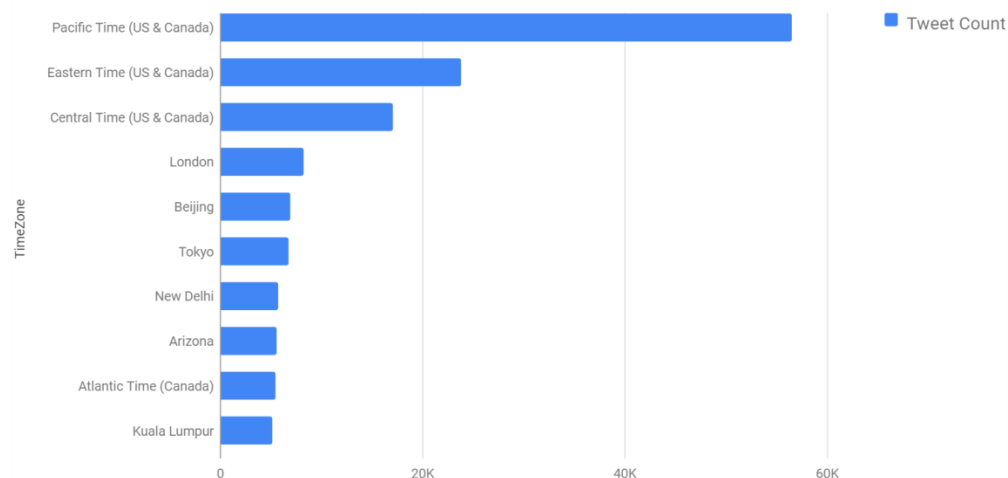


***More number of tweets are tweeted from which time zone***

```
SELECT user.time_zone, count(*) as c  
FROM EntertainmentTable WHERE user.time_zone <> 'null' GROUP BY user.time_zone  
ORDER BY c desc limit 10
```

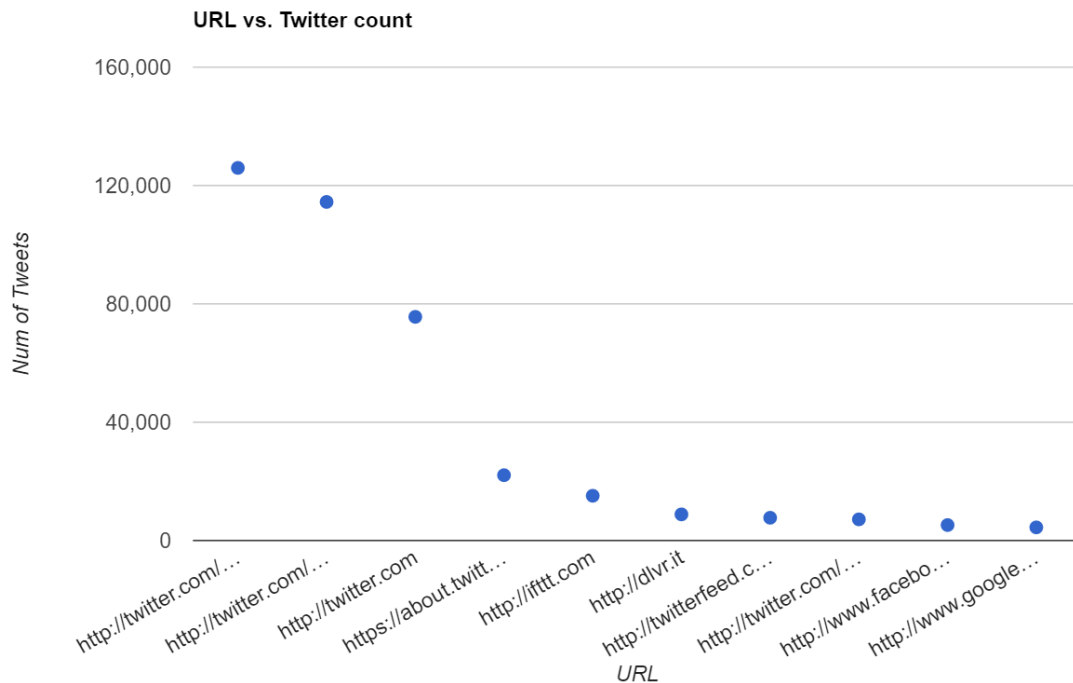
TimeZone and Tweets

We know from where you tweeted



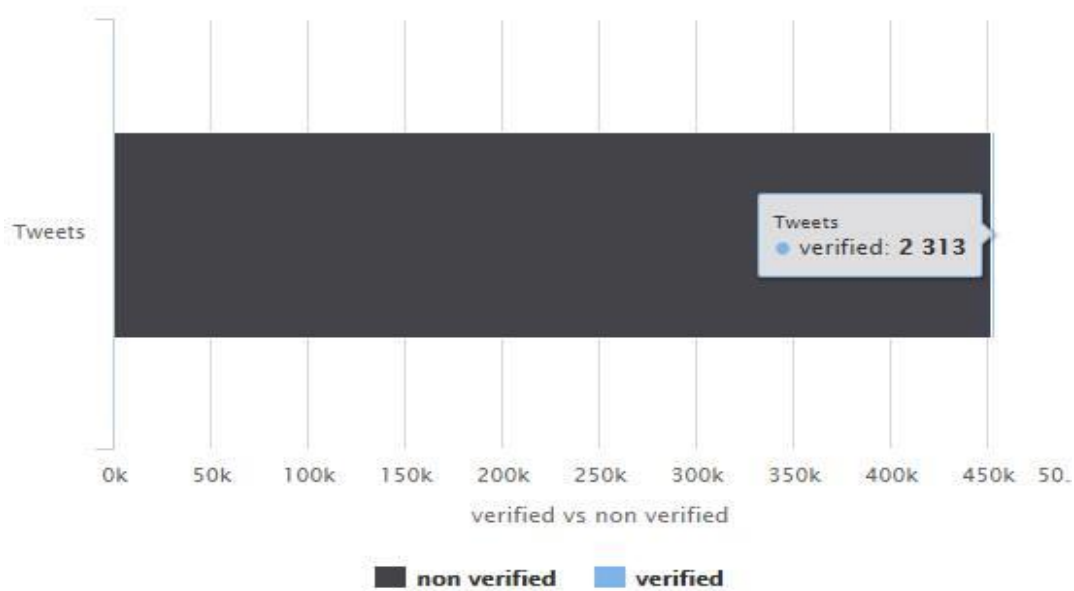
***What are the top most popular URL links used in the tweets?***

```
SELECT source, count(*) as c FROM EntertainmentTable WHERE source <> 'null' GROUP BY
source ORDER BY c desc limit 10
```



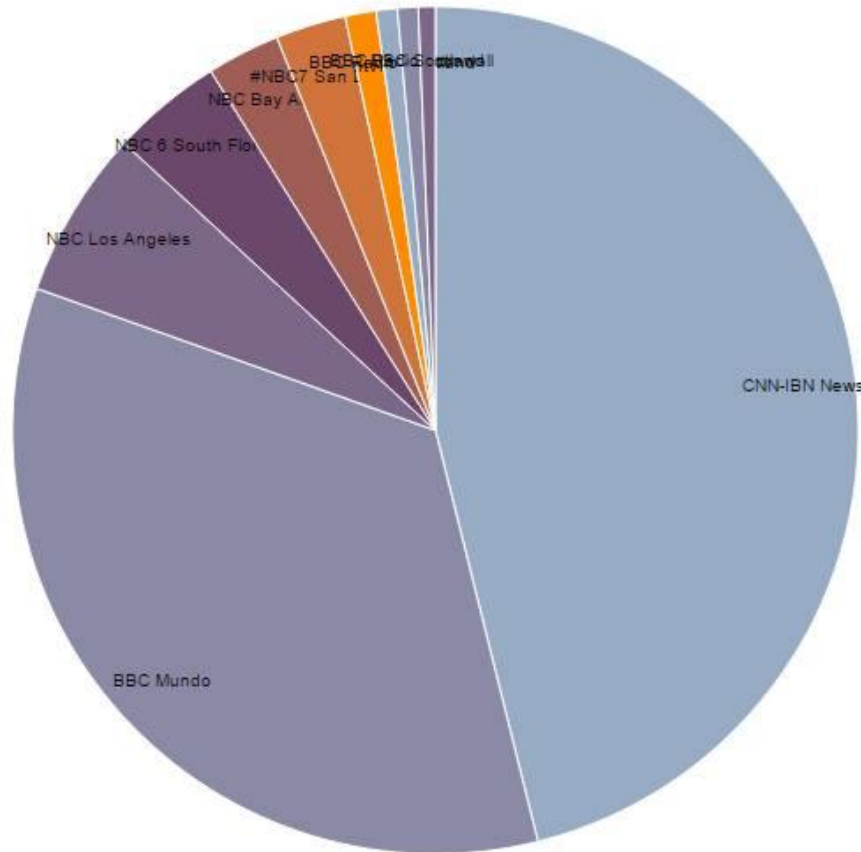
***Users who are verified and non-verified***

```
SELECT if(user.verified=true,'verified_users','nonverified_users') as status,count(*) FROM
tweets4 GROUP BY user.verified LIMIT 2
```



***Top most new channel followers***

```
SELECT user.name,max(user.followers_count) as c FROM tweets2 WHERE user.name like '%NBC%' or user.name like '%CNN-IBN%' or user.name like '%fox%' or user.name like '%BBC%' or user.name like '%tv9%' or user.name like '%ndtv%' or user.name like '%ntv%' GROUP BY user.name ORDER BY c desc limit 10
```



**Visualizations:**

1. Bar Diagram
2. Pie Chart
3. Representing on Maps
4. Donut Diagram
5. Scatter Chart
6. Stringy
7. Stacked bar

**References:**

<http://spark.apache.org/docs/latest/sql-programming-guide.html#json-datasets>

<https://dev.twitter.com/overview/api/tweets>

<https://github.com/dhotson/springy/>

<https://developers.google.com/chart/interactive/docs/gallery/map#customizing-markers>

<http://www.highcharts.com/demo/pie-gradient>

<http://bl.ocks.org/mbostock/3887235>

<https://developers.google.com/chart/interactive/docs/gallery/barchart#overview>

<https://developers.google.com/chart/interactive/docs/gallery/piechart#making-a-donut-chart>

<https://developers.google.com/chart/interactive/docs/gallery/scatterchart>

<http://www.highcharts.com/demo/bar-stacked>