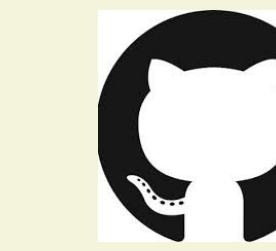


# PRINCIPLES OF BIG DATA MANAGEMENT PROJECT

Department of Computer Science Electrical Engineering, University of Missouri-Kansas City, Missouri



Mehmet Acikgoz



acikgozmehmet/Twitter\_Analysis

## INTRODUCTION

In this project we design, develop and deploy an application that performs and visualizes some sophisticated analytic queries executed on twitter data, which are related to Turkey's military action on the northern part of Syria on October 14, 2019.

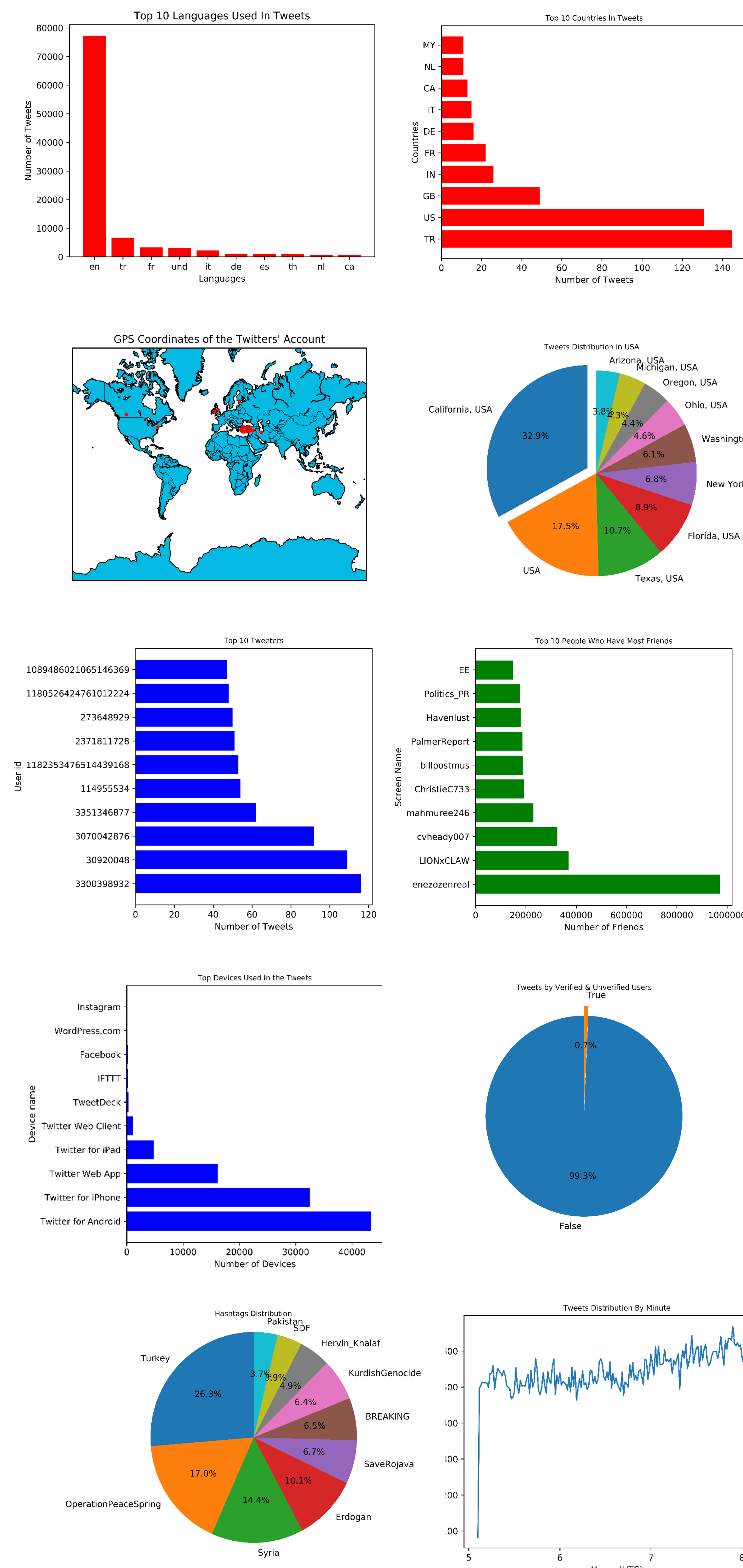
## OBJECTIVE

The main objectives are to collect, store and analyse the twitter data by using state-of-art technologies in Big Data and to visualize the results by making use of some powerful visualization tools.

## TECHNOLOGIES USED



## ANALYSIS



## METHODS

We collected 100.000 tweets filtering twitter data by #Recep Tayyip Erdogan, #Erdogan, #Turkey, #PeaceSpring, and #OperationPeaceSpring. Then, we stored them in a file to use as the input file for the analysis in the subsequent step.

During the analysis, we used PySpark to run 10 complex queries, and matplotlib to visualize the results.

We also created a webpage with HTML & CSS to present the results locally.

Last but not least, we containerized the application with Docker so that it can be shipped and deployed to any platform easily.

## CONCLUSION

We have gained invaluable hands-on experience in not only using Big Data technologies, but also understanding them for data analysis and visualization.

We have gotten insights about Turkey's military activity on the northern part of Syria by analysis and visualization of tweets.

**Under the Guidance**  
**Dr. Praveen RAO**

Department of Computer Science Electrical Engineering  
University of Missouri-Kansas City