Identifying Trends on Twitter with Apache Spark and Python   
\*(I would like to create a Twitter data pipeline using Spark and Python to analyze hashtags and identify trends. I'm building this as a project for my GitHub in order to boost my data engineering job applications and brush up on my Spark.)

* Build simple application that reads online streams from Twitter using Python
* Process tweets using Apache Spark Streaming to identify hashtags
* Return top trending hashtags (represent data on a real-time dashboard)

1. I need an Access Token *{An object that identifies and authenticates the user/process/thread and their privileges}* from Twitter (<https://developer.twitter.com/en/apps>).
   1. **API key**: jnTEuu5nSRGwXyEHbLBjvgzHw
   2. **API secret key**: DKahlLSTUpS3jvgq8FGQ8E5rpJ22bY7upKOQM364TIU2O8IUlE
   3. **Bearer token** (*Access token* used with OAuth 2.0): AAAAAAAAAAAAAAAAAAAAAMo9GQEAAAAA%2ByNoo05ShDWmP4p%2Bp52H99RfuJg%3DimJxqfvLEoZbHlK2hiLFboN4d21nmMH7xbtIzEg8tEQuXOGYd0
   4. **Access token**: 37060595-FlShwh35livHmS7hozzhUSCViWEhJ4C7feLpdtBjW
   5. **Access token secret**: AwcydUYWbIhDBNoChahtyZ0u7d97NGA5eCsTvWem1WL7B
2. Building the Twitter HTTP Client
   1. Build client to from Twitter API using Python
   2. After, pass them to Spark Instance
3. Write Python Script for Server Socket and Client Socket Connection (I tested Twitter feed from server to client)
   1. Configure IP Address to be localhost and port 9009
   2. Call get\_tweets and pass response with socket connection to send\_tweets\_to\_spark to send to Spark
4. Setting up Apache Spark Streaming Application (for Real-Time Processing of incoming Tweets)
   1. Extract Hashtag from Tweet
   2. Calculate number of hashtags