PROBLEM 1

Data source:
I am collecting the Twitter data using Twitter's Streaming API calling it through streamR package as provided by R. The authentication is being does using OAuth 1.0 specification.

Data Structure:
The data is in the standard JSON format.

Duration of data collected:
I have collected the data over a period of 7 days from Feb 22nd till March 4th.

Packages used:

A. streamR - The package provides a series of functions that allow us to access Twitter's filter, sample, and user streams, and to parse the output into data frames. For this problem I am making use of the two of these functions:
   - filterStream - opens a connection to Twitter's Streaming API that will return public statuses that match one or more filter predicates. Tweets can be filtered by keywords, users, language, and location. The output is being saved into a JSON file.
   - parseTweets - converts tweets in JSON format to data frame.

B. RCurl - Provides functions to allow one to compose general HTTP requests and provides convenient functions to fetch URIs, get & post forms, etc. and process the results returned by the Web server. We are using it to connect to Twitter API.

C. ROAuth - Provides an interface to the OAuth 1.0 specification allowing us to authenticate via OAuth to Twitter using the secret public and private keys. The parameters passed to the streaming endpoint are as follows:
   - requestURL
   - accessURL
   - authURL
   - consumerKey
   - consumerSecret

References:
https://cran.r-project.org/web/packages/