Steps involved:

1. Used Twitter Streaming API to download tweets related to movie PK.
2. Getting twitter API keys: In order to access Twitter Streaming API, we need to get 4 pieces of information from Twitter: API key, API secret, Access token and Access token secret.
3. Connecting to twitter streaming API and downloading data by using “tweepy”
4. Program run for approximately 12 hours to collect about 100 mb data.
5. Download 4 Python libraries json for parsing the data, pandas for data manipulation, matplotlib for creating charts, and re for regular expressions.
6. Clean and structure the extracted data.
7. Analysing the trends in tweets according to language used, country and sentiment present in the tweets for the topic i.e. movie pk.
8. Used tagged words with scores from <https://github.com/uwescience/datasci_course_materials/blob/master/assignment1/AFINN-111.txt> for doing the sentiment analysis using the bag of words model by finding the sentiment carrying words in the tweets using regex and hence giving the sentiment value as positive or negative (happy or sad).
9. Report generation by using pie charts and barcharts.