Spatial Temporal Tweet Analysis

Tejal Patted   **Fall 2017**

Introduction

Micro blogging websites such as Twitter, Tumblr have evolved to become great source of opinion forum. Sentiments in tweets can be leveraged to understand variation in people's emotion towards varied subjects, during various times of the day and at different locations. This project analyzes sentiment variation at different types of locations using twitter data over entire USA. Further on we explore hourly sentiment variation from Thursday through Saturday.

Hypothesis

Spatially we expect to find significantly lower concentration of negative tweets around universities, parks and airports which indicates positive emotions. Further on, locations such as hospitals and police stations are expected to have higher concentration of negative tweets. For the temporal analysis, we assume to find higher percentage of negative tweets on Thursday and a gradual decline in percentage by Friday evening.

Methodology

Twitter data was scarped using REST API and python *tweepy* library. Sentiment analysis was achieved using python's *NLTK* library. Spatial Analysis was done using custom model built in *ArcGIS* model builder. Temporal analysis was done using *Matplotlib* library to create visualization
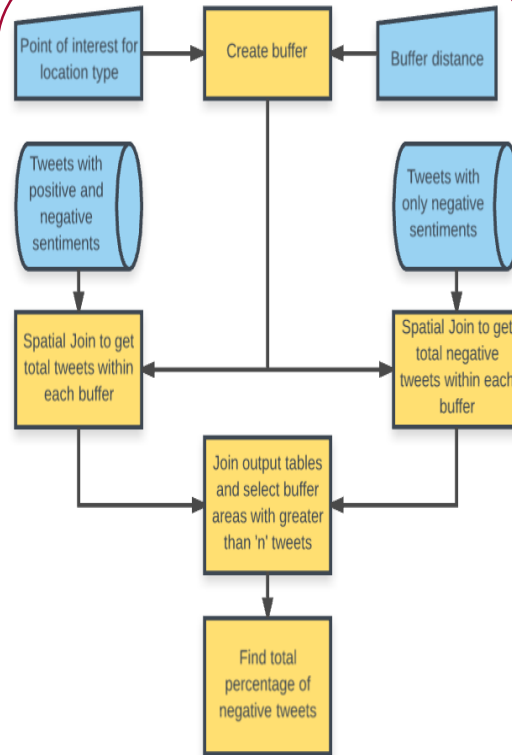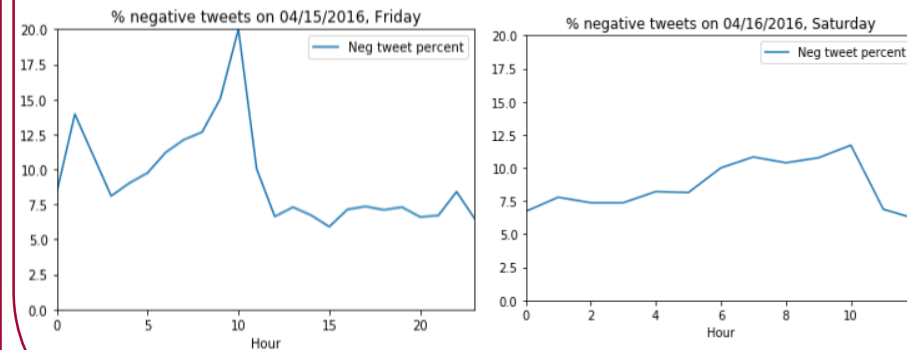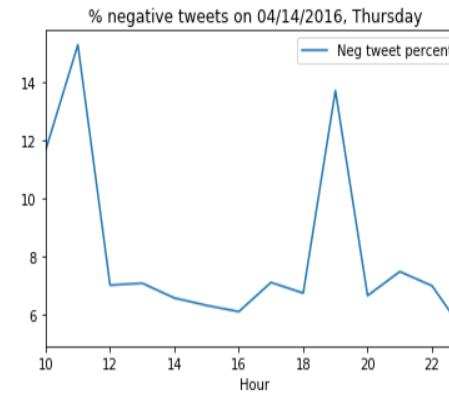

Fig1: Spatial analysis custom model

Fig2: Temporal analysis visualization

Spatial Model & Analysis

Each point of interest such as university, hospital, parks were considered separately with 1 mile buffer radius. The universities, parks and airports have low concentration of negative tweets between 5-8%. The hospital had 5.6% of tweets to be negative which cannot prove our hypothesis.

Conclusion

Airports, Universities and parks had low negative tweet concentration. However, no locations were found to have significant concentration of negative tweets. This could be due to skewed twitter data that had only 19% negative tweets. Negative tweets were high at 10 am and 7 pm on Thursday and 10 am on Friday which correspond to traffic hours. The negative tweets were considerably low starting Friday evening and remained low through out Saturday