CKME 136 Data Analytics: Capstone

Prepared by: Chris Kogut

Project Abstract

1. Theme for my project will be Sentiment Analysis based on the Twitter Data.

Coffee is one of the most popular drinks in Canada. People order coffee from various vendors based on price, taste, location.

My goal is to evaluate coffee preferences from two companies Tim Horton’s and Starbucks.

I’ll try to answer questions that would allow me to select best company and best coffee.

To solve the “puzzle” I’ll try to find answer to the questions defined below:

Which company serves best coffee?

Which region prefers which vendor?

Optional:

What is the most popular coffee type from each vendor?

What is the most and least popular time to drink coffee:

* Day(Morning, afternoon, evening)
* Weekday, which day is most least popular for coffee drinkers

Analysis will be performed based on the twitter data, and I will using HashTags associated with each company. I will be searching for a specific keywords that would describe coffee and associated sentiment. In my work, I will be collecting these sentiments and classify polarity of sentiments in these opinions w.r.t. Positive, Negative or Neutral.

Twitter data will be collected for analysis using Twitter API.

Approach: I’ll be using Dictionary Based approach to analyze data posted by different users. Then polarity classification of this data will done.

R will be used to select and evaluate the data to answer these questions.