**EECS4415 – RoadMap**

**Data Collection**

> Collect data through Twint

* + Split by provinces -> Ontario, Quebec etc...
  + Split by Main City's -> Toronto, Vancouver, Quebec City
  + Split by ridings -> North York, Downtown etc...

> Split data for each party; Liberal, Conservative, NDP

**Cleaning Data**

>Remove Emoji’s, URL links, Non-Char’s, Translate French

**Analysis/Tasks**

1. **What is the political sentiment distribution of major geo-locations in Canada?**
   * 1. Run a sentiment analysis on the provinces, cities and ridings
     2. Run it for each Liberals, Conservatives, NDP
     3. **Goal:** Create a chart of sentiment vs time (1 month intervals) that shows the distributions sentiment of each party's
2. **Which riding/city/province have a positive/negative sentiment towards a political party?**
   * 1. Using the sentiment analysis data, create a map that shows which party is the most favored (most positive) for each province then for each city and lastly each riding
     2. **Goal:** Create a Map showing the most positive party
3. **What do people value in each riding/swing state? (e.g. Taxes, housing, education, etc.)**
   * 1. Using our data set of tweets, look for the occurrence of keywords in each tweet for each province/city/riding
     2. TOPICS: Crime, Education, Environment, Families, Foreign Policy, Government, HealthCare, Immigration, Jobs & Economy, Science & Tech, Taxes
     3. For each topic create a set of keywords. Use those keywords and total the sum of each of them to create a total value for each topic
     4. **Goal:** Create a bar chart that shows the value for each topic of each location
4. **What’s the percentage that a specific riding will have a political shift in the upcoming election?**
   * 1. Figure out the ridings that change the most (ridings with 50/50 sentiment)
     2. Ridings that are closer to 50/50 sentiment split have a higher percentage that it will have a political shift
     3. **Goal:** Create a map that shows the percentage of each riding in terms of whether it is closer to 50/50 or not
5. **How does the public sentiment about a candidate change over time?**
   * 1. Collect new data that is only about the candidate for each party (Use twint and search only for Trudeau, Scheer, and Jagmeet)
     2. Run a sentiment analysis on that data
     3. **Goal:** Determine two things;
        1. The current sentiment for each political figure
           1. Create a circle chart of positive, negative and neutral sentiment
        2. The sentiment over time of 1-month intervals
           1. Create a sentiment vs time (1-month intervals) of each candidate
6. **Streaming (Real Time Sentiment)**
   * 1. Collect real time political tweets
     2. As they are collected run a sentiment analysis on the tweets
     3. **Goal:** Create a real time sentiment vs time chart for each party (interval every 1 min?)

**End Goal:** Create a visual and interactive analysis platform

