Proposal

Target Group

Government of Canada - Sevice Canada, The Canada Border Services Agency (CBSA)

Dataset

- Getting the data: Acquiring/gathering/downloading.
    - https://www.ncei.noaa.gov/products/land-based-station/global-historical-climatology-network-daily

Problem Statement

Canada has always been a safe haven for refugees during war and crisis. In light of recent events, more than 32000 Ukrainians have found support and a safe place to live here. While Canada is a large country, most of it is not very suitable for living, especially for those seeking refuge. Our aim for this project is to provide data backed study of different locations in Canada where the government could expand cities and build suitable shelters. The factors to be considered would be the weather of the region including temperature, rainfall, snowfall, and snow depth and the occurrence of disasters like cyclone, tornado, and blizzards. With this information, the government could spend resources efficiently in setting up transportation, crops and livestock, infrastructure and housing, and sustainable energy sources.

Workflow

- Algorithmic work: Work on the algorithms including integrating data mining and machine learning techniques.
    - ETL
    - Data analysis and evaluation
    - Data visualization
    - Bigness/parallelization

Technologies

- ETL: Extract-Transform-Load work and cleaning the data set.
    - Spark
- Bigness/parallelization: Efficiency of the analysis on a cluster, and scalability to larger data sets.
    - AWS
- Visualization: Visualization of analysis results
    - Tablue or Matplotlib or seaborn or Kibana
- Technologies: New technologies learned as part of doing the project.
    - AWS (expanding AWS usage beyond the scope of the assignments)
    - GitHub CI/CD (actions)
    - Tablue or Matplotlib or seaborn or Kibana

Team Collaboration

- Members: Karishma, Crystal, Jeanne
- Version Control System: SFU GitHub