**Project Proposal: Predicting Immigration Trends to Canada in the Next Five Years**

**Course:** CS 2704  
**Date:** March 29, 2025  
**Team Members:** as discussed no group yet preferably individual project  
**GitHub Repository:** https://github.com/fmsl11/cs2703.git

**1. Project Title**

**Forecasting Immigration Trends to Canada (2025-2030) Using Data Analytics**

**2. Introduction and Background**

Canada has long been a leading destination for immigrants due to its strong economy, favorable immigration policies, and high quality of life. Immigration plays a crucial role in Canada's labor market and demographic growth. In recent years, government policies have aimed at increasing immigration targets to address labor shortages and an aging population. This project aims to analyze historical immigration trends and predict the number of immigrants to Canada over the next five years using statistical modeling and machine learning techniques.

**3. Hypothesis**

We hypothesize that the number of immigrants to Canada will continue to increase over the next five years due to favorable government policies, economic opportunities, and global migration trends. We expect our analysis to reveal a consistent upward trajectory in immigration numbers.

**4. Dataset Selection**

* **Dataset Name:** Canadian Immigration Data (1990-2025)
* **Source:**
  + Government of Canada Open Data Portal (<https://open.canada.ca/data>)
  + United Nations Migration Reports (<https://www.un.org/en/development/desa/population/migration>)
  + World Bank Immigration Statistics (<https://data.worldbank.org/indicator/SM.POP.NETM>)
* **Description:** The dataset includes yearly immigration numbers, breakdowns by country of origin, economic class, refugee status, and employment sectors. Additional variables such as GDP growth, unemployment rate, and policy changes will also be considered.

**5. Plan for Testing the Hypothesis**

1. **Data Cleaning and Preprocessing:**
   * Handle missing values, normalize data, and convert categorical values where necessary.
   * Merge immigration data with economic indicators and policy changes.
2. **Descriptive Analytics:**
   * Visualize trends in immigration over the past three decades.
   * Compute basic statistics such as mean, median, and standard deviation for annual immigration rates.
3. **Predictive Analytics:**
   * Use linear regression, time series forecasting (ARIMA), and machine learning models (Random Forest, LSTM) to predict immigration numbers for 2025-2030.
   * Compare model accuracy using RMSE and R-squared metrics.
4. **Visualization and Interpretation:**
   * Generate line graphs, correlation heatmaps, and predictive trend plots.
   * Interpret key insights regarding policy impacts and economic drivers of immigration.

**6. Expected Outcome**

* We anticipate that immigration to Canada will show a steady increase, influenced by government policies targeting labor shortages and demographic needs.
* Predictive models will indicate a positive correlation between economic growth and higher immigration rates.
* Possible deviations may occur due to external factors such as global conflicts, economic recessions, or sudden policy shifts.

**7. Materials for Submission**

* **Slides explaining the hypothesis and methodology.**
* **Dataset snippet or link to the source.**
* **Code repository for data processing, analysis, and visualization.**

This proposal sets the foundation for a comprehensive study on Canadian immigration trends, integrating data-driven insights with policy and economic analysis.