**Project Overview & Instruction**

## **Project Overview**

Project Name: Increasing Public Data Transparency for Immigration Law in Canada

#### **Description of the problem/question:**

**Part 1**: IRCC has provided us with data about inadmissibility refusal applications, which need to be examined for trends, patterns, and other information that are best understood from a data science lens, rather than an immigration-focused lens.

**Part 2**: In parallel, we’d like to replicate the study done by Professor Sean Rehaag in [Luck of the Draw III: Using AI to Extract Data About Decision-Making in Federal Court Stays of Removal](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4322881) to supplement part 1. This will allow us to extract additional and complementary insights on the data provided by IRCC, by exploring the appeal cases and judicial reviews.

The professor encourages law professionals and academia to use this data for non-profit purposes with proper citation. The detail of customization is noted in the instruction below.

We intend to conduct an exploratory data analysis to gain relevant summary statistics as well as insights into associations between numerous variables in the dataset, which could result in relevant hypotheses for further exploration.

Finally, the results of these models should be presented in a dashboard or PPT format for non-technical readers.

We are also leading advocates around data transparency. IRCC has been gatekeeping data, releasing it selectively to those who request it and have not set out clear perimeters. Meanwhile, applicants who are seeking to litigate their cases and show statistical trends are unable to provide this data. By publishing a dashboard with Government data, we are pushing for more data transparency and clarity in line with IRCC’s own openness, transparency, and accountability commitments: www.canada.ca/en/immigration-refugees-citizenship/corporate/transparency.html

#### **Problem/question impact:**

As immigration professionals, our day-to-day work is focused on serving client needs by answering legal questions, filing applications, and litigating unreasonable decisions. We are also stakeholders in IRCC’s work as their internal processing rules and policy and program changes directly impact our clients; however, the data is often not publicly available and when it is disclosed, the data received is unprocessed. We require assistance to process datasets obtained from IRCC and analyze federal court cases to help us understand how we should navigate future files and help similarly situated practitioners to better serve their clients.

This is also occurring as IRCC is moving to Advanced Analytics-based decision-making that is transforming the process of immigration and requires a better understanding of the data challenges.

Our non-profit advocacy work through AIMICI, would benefit from being able to draft/write reports based on this data dashboard and to share with stakeholders (other lawyers, Government, judges) to transform the work in this space.

## **Legal Instruction**

This project will focus on the legal topic of “inadmissibility”. Here is the introduction of the topic for you to familiarize yourself with terms.

**What is Inadmissibility?**

Inadmissibility refers to a legal finding that a foreign national or permanent resident is not allowed to enter or remain in Canada under the Immigration and Refugee Protection Act (IRPA).

**Key Points:**

* Individuals found inadmissible may be denied entry, refused a visa, or subject to removal from Canada.
* Inadmissibility can apply at various stages: application for a visa, at the border, or during a review of status in Canada.
* It is assessed by immigration officers or the Immigration Division (ID) of the Immigration and Refugee Board (IRB), depending on the case.
* May lead to refusal of entry, removal orders, or loss of immigration status.

**Inadmissibility Type**

| **Ground of Inadmissibility** | **IRPA Section** | **Explanation** | **Leading Cases** |
| --- | --- | --- | --- |
| Security | s.34 | Involvement in espionage, terrorism, subversion (attempts to overthrow a government, etc.), or threat to Canada's security. | *Agraira v. Canada, Suresh v. Canada* |
| Human/International Rights Violations | s.35 | Participation in war crimes, crimes against humanity, or being a senior official in a regime engaged in such violations. | *Ezokola v. Canada, Mugesera v. Canada* |
| Serious Criminality | s.36(1) | Conviction with a sentence of ≥6 months in Canada or an equivalent foreign offense. | *Tran v. Canada, Medovarski v. Canada* |
| Criminality (less serious) | s.36(2) | Less serious crimes or summary convictions in Canada or abroad. | *Williams v. Canada, Hill v. Canada* |
| Organized Criminality | s.37 | Membership or involvement in criminal organizations or transnational crime. | *Medovarski v. Canada, Agraira v. Canada* |
| Health Grounds | s.38 | Medical conditions that may endanger public health/safety or cause excessive demand on health or social services. | *Hilewitz v. Canada, De Jong v. Canada* |
| Financial Reasons | s.39 | Inability to financially support oneself, potentially relying on social assistance. | N/A |
| Misrepresentation | s.40 | Providing false information, withholding facts, or misrepresentation in an immigration process. | *Sayedi v. Canada, Baro v. Canada* |
| Non-Compliance with IRPA | s.41 | Breach of immigration requirements (e.g., overstaying, working without authorization). | N/A |
| Inadmissible Family Member | s.42 | Family member is inadmissible under any other ground; affects applicant’s admissibility. | N/A |

**Other keywords:**

**Keywords**

| **Ground of Inadmissibility** | **IRPA Section** | **Keywords** |
| --- | --- | --- |
| Security | s.34 | Espionage  Against Canada  Canada’s Interests  Subversion  Democratic Government  Terrorism  Danger to Security  Violence  Endangerment  Membership  Complicity |
| Human/International Rights Violations | s.35 | Human Rights  International Rights  Violations  Senior Official  Government  Regime  Terrorism  Genocide  War Crimes  Crimes Against Humanity  Participation  Contribution |
| Serious Criminality | s.36(1) | Criminal Conviction  Foreign Conviction  Imprisonment  10 Years/Ten Years  Sentence  Over 6 months/six months  Serious Indictable Offence  Commission |
| Criminality (less serious) | s.36(2) | Criminal Conviction  Foreign Conviction  Indictment  Indictable Offence  Summary Offences  Commission |
| Organized Criminality | s.37 | Membership  Criminal Activity  Organized Crime  Acting in Concert  People Smuggling  Trafficking  Money Laundering  Proceeds of Crime |
| Health Grounds | s.38 | Danger to Public Health  Danger to Public Safety  Excessive Demand on Health Services  Excessive Demand on Social Services |
| Financial Reasons | s.39 | Unable or unwilling to support oneself  Unable or unwilling to support dependents  Arrangements for care and support  Social Assistance |
| Misrepresentation | s.40 | Misrepresenting  Withholding  Material Facts  Error in Administration  Non-disclosure  Omission  False Statement  False Information |
| Non-Compliance with IRPA | s.41 | Contravention  Non-Compliance  Failure to Comply |
| Inadmissible Family Member | s.42 | Inadmissible Family Member  Accompanying Family Member |

\*“Reasonable grounds to believe” is a keyword for ss. 34 to 37.

**Exclusion keywords:**

For this research, we would like to exclude refugee claims. Please use the keywords below.

* Convention refugee
* Person in need of protection
* Refugee claimant
* Protected person

**Keywords for decisions:**

**Orders of Appeal after Inadmissibility Findings**

When an individual is found inadmissible to Canada, they may have access to various avenues of recourse depending on their immigration status and the grounds of inadmissibility. The process typically begins with a hearing before the Immigration Division (ID) to assess the validity of the inadmissibility finding. If applicable, certain applicants may appeal to the Immigration Appeal Division (IAD). Others may seek judicial review at the Federal Court. In rare cases, further appeals may proceed to the Federal Court of Appeal and the Supreme Court of Canada if leave is granted.

|  | Where | Data Source | Note |
| --- | --- | --- | --- |
| 1 | Immigration Division (ID) | Excel Data |  |
| 2 | Immigration Appeal Division | Not Available | Not all applications / applicants are allowed to appeal at IAD. |
| 3 | Judicial Review at Federal Court | FC |  |
| 4 | Federal Court of Appeal | FCA | Only if the Federal Court certifies a serious question of general importance. |
| 5 | Supreme Court of Canada | SCC | Discretionary and rarely granted. |

\*\*Data source name is corresponding to <https://huggingface.co/datasets/refugee-law-lab/canadian-legal-data>

## **Technical Instructions**

This project with the Heron Law Office (AIMICI) has two parts and first we will explain the legal terms for students to be familiar with this topic.

**Part 1: Exploratory Analysis**

We have excel files of the Immigration application data related to inadmissibility from the IRCC (Immigration, Refugee and Citizenship Canada).

Data source (example): <https://docs.google.com/spreadsheets/d/136rLvjSNkfuELQDQtwqYvxLUHYtrQoy2/edit?gid=1753241632#gid=1753241632>

We’re currently requesting the same data in a tabular format from the IRCC on all inadmissibility grounds.

**Part 2: Case Law**

This part of the exercise leverages the research conducted by the York University’s Professor Sean Rehaag’s code.

Github: [Luck of the Draw Code](https://github.com/Refugee-Law-Lab/luck-of-the-draw-iii/tree/main)

Dataset: [Supreme Court of Canada Bulk Decisions](https://refugeelab.ca/bulk-data/scc/)- data is scraped from [Federal Court site](https://www.fct-cf.gc.ca/en/court-files-and-decisions/court-files)

In the original research, the topic focused on understanding the decision making in the “stay of removal” cases.

For our research, we’d like to focus on “Inadmissibility” cases explained in the Legal Introduction section.

We’re interested in the analysis and findings by inadmissibility type in the table above and you can use keywords listed in the table above.

Datasource**:** <https://huggingface.co/datasets/refugee-law-lab/canadian-legal-data>

<https://refugeelab.ca/bulk-data/scc/>

Questions to explore:

* Are there biases in the decisions? judge, citizenship, location or certain attributes
* Any trends in the time series? recent trends vs. past trends
* What are other insights found in the data?

**Caution: Deviation from the original Code**

A few things to note when you’re running the professor’s original codes.

1. Error in the code

When you run the professor’s codes (1~6) in JupyterNotebook, you’ll encounter a few errors.

Example: you’ll need to run “Read Me” code first before running #1-6. You will need to specify the location of saving the parquet file in code 1. You’ll need to replace OpenAI API keys and model names. I will share the list of these in the coming weeks as much as possible.

1. Version of Open AI Model

The original code uses the “gpt-3.5-turbo”. It is ideal to switch with the “gpt-4o-mini” for cost performance.

1. Training Files

The original code uses three training human coded excel files below.

https://github.com/Refugee-Law-Lab/luck-of-the-draw-iii/tree/main/training\_data

Instead of using excel files, we would like your team to create the training data sets for categorizing cases using keywords from the legal section.

You may be able to use the same file for extracting judge names as this is more generic.

I'm also testing part 2 of the professor's original code, so I will share my findings in the coming week. As part 2 may require some code changes, your team can come up with alternative models or solutions as well.

1. Share final GitHub codes

Please publish your entire codes to GitHub and share with us at the end of