INACCESSIBLE JUSTICE

A qualitative and quantitative analysis into the Demographics, Socioeconomics, and Experiences of Self-Represented Litigants

David Lundgren

Table of Contents

1. Introduction

The Department of Justice ensures that the Canadian justice system is as "fair, accessible, and efficient as possible" (Department of Justice, 2018). It seeks to provide equitable access to justice for all persons in Canada to promote respect for the law and the rights and freedoms of citizens in the absence of structural barriers and undue discrimination. These principles uphold equality under the law and independence for persons to seek legal assistance and redress in a system designed to instil and enshrine justice. At the very least, it is supposed to.

Inconsistency and prejudice in the justice system continue to plague agents looking to the courts to resolve legal problems that burden everyday experiences. No less so does it seethe through the barracks of those representing themselves. Self-represented litigants often find themselves fighting on two fronts: having to deal with the process to resolve their legal claim while at the same time battling the inefficiencies and misperceptions of self-representation. Those who represent themselves are at an inherent disadvantage, through no fault of their own, by having to familiarise themselves with the legal structures and processes that opposing lawyers have dedicated their lives to understanding. While the right to self-represent constitutes part of the foundational principles that underlie the court system and is afforded to each citizen, its mark is fundamentally at odds with the courts' mission and ideal of distributive justice.

This paper centres around a more realistic characterization of who self-represented litigants are and the issues they face. For various social, economic, or geographic reasons, self-represented litigants tend to have unmet legal needs, increasing the cost of already burdensome and cumbersome judicial proceedings. These result from an overall lack of legal resources and assistance, low incomes, low education, and low digital literacy rates, often leading to misunderstandings of social and legal needs and court processes. The disadvantaged position of those self-representing leads to power imbalances in the courtroom that reduces their access to justice. Further, the conflation between behaviours deemed to be vexatious and behaviours

that suggest confusion and unintended errors that may result from unfamiliarity with courtroom proceedings is concerning.

The amalgamation of these factors has broad and far-reaching consequences that affect more than just the individual person. Misperceptions of self-represented litigants and power imbalances in the courtroom lead to more lengthy proceedings, a proliferation of legal cases, and higher court costs. In a justice system intended to promote the ideal of efficiency, it is least realized in the case of self-represented litigants. The spill-over costs lead to an undersupply of effective legal services, misallocation of resources, and inherent issues with procedural fairness. These issues manifest in economic losses, which plague the entire court system and begs the question of whether the ideal of justice is realized in practice. In short, this paper puts forth that the nature and misperception of self-represented litigants have induced inefficiencies in the justice system that impose barriers on their access to justice.

This paper is structured into two main parts. The first part contextualizes the issue and provides an overview of Canada's current state of self-representation. Notably, this section explains how the reality of self-representation starkly contrasts the ideal purported by government sources. It also addresses 'Who are self-represented litigants?' and how they are treated based on sex, socio-economic factors, and vexatious designations. Overall, it is clear that self-represented litigants suffer from a myriad of problems that are inherently connected to barriers to accessing justice. The second part introduces the data used in this analysis, taken from the National Self-Represented Litigants Project's annual Intake Form. It provides a brief description of the methodology used in the analysis portion and an introduction to machine learning models to contextualise the discussion. The results examine the intersections between the self-represented problem and social injustices stratified along gender, ethnicity, income, age, and education levels.

2. The Ideal versus Reality of Self-Representation in Canada

2.1 The Ideal of Self-Representation

A divergence from the ideal of fair and equal access to justice characterises Canada's current state of self-representation. However, there is a lack of awareness of this injustice as the idyllic and ambitious principles of self-representation that guide the courts subdue any insufficient or inefficient findings in practice to the public.

The government and Department of Justice enshrine the ideal of self-representation in its *Statement of Principles on Self-Represented Litigants and Accused Persons* (Canadian Judicial Council, 2006). The Canadian Judicial Council predicates these principles on the responsibility to ensure that fair access and equal treatment by the courts are available to all self-represented litigants. However, they note that unrepresented litigants face, and often present themselves, systemic challenges in their interactions with the courts (p. 3). Therefore, a common thread throughout the Council's report is that persons "seeking access to the court should be represented by counsel" (p. 1). This inherently delineates the importance of the principles wherein many persons seeking access to the courts do not have the opportunity to be represented. It is of concern that such remarks are made in the document meant to tackle those issues. Nevertheless, the Council proceeds to provide guidance and ideals to litigants and judges across three overarching categories: (1) equal access to justice, (2) promoting a thorough understanding of the field of law, and (3) the fundamental disadvantages of those self-represented.

The right to access and equal justice derives from the constitutional pillars of our justice system. Concerning self-represented litigants, as noted by the Council, this right may be strikingly contentious (p. 2). The *Statement* details that the average person "may be overwhelmed by the simplest of court procedures", and more so if they face stress and external

challenges outside the courtroom (p. 3). The confusion and complexity of legal procedures can intimidate the average person whose lack of knowledge or skill may affect their ability to actively participate in and litigate their case (Hann et al., 2002, p. 20). In response, the Council affirms in the *Statement* that judges should engage affirmative and non-prejudicial approaches with self-represented parties (Canadian Judicial Council, 2006, p. 5). However, this creates a tension between a judge's discretion to accommodate self-represented litigants and the requirements of judicial impartiality and neutrality as outlined in the Ethical Principles for Judges (Canadian Judicial Council & McLachlin, 2004).

Further, adding to this tension are judicial decisions outlining that equal treatment does not result in equal justice, and that it is for the judge to "rectify and prevent" discrimination against particular groups (*Eldridge v. British Columbia (Attorney General),* 1997, para. 54*).* While this tension allows for broad variation in judicial proceedings and discretion, the Council offers no practical guidance apart from the idyllic principles outlined in the *Statement.* A lack of effectual standards and roadmaps for judges dealing with self-represented litigants exacerbates their marginalisation in the legal system. It renders such promises of equal treatment and fair access to justice unrealisable.

The *Statement* details that self-represented litigants need to develop the correct understanding of and appreciation for legal formalities and structure in their case. The Council writes that the realisation of the ideals of justice for self-represented litigants is conditioned on their "awareness and understanding of both procedural and substantive law" (Canadian Judicial Council, 2006, p. 1). While the *Statement* offers no guidance as to the extent of awareness and understanding necessary, it makes clear that self-represented litigants are at an inherent disadvantage in meeting this threshold. Not only do they have to balance their private, professional, marital, and parental obligations, but they also must develop an understanding of the law, which many lawyers rely on years of study to comprehend. The Council does note that

leniency for minor deficiencies is to be tolerated, but what amounts to a minor deficiency is left unclear and to the judge's discretion (p. 7). It is concerning that a minor deficiency in the eyes of a self-represented litigant is likely inconsistent with what a judge would rule as a minor deficiency, who presumably has developed a much higher standard for what is acceptable through years of experience. Thus, the vague requirement that self-represented litigants need to achieve some understanding threshold provides little help or assistance to those looking to the Council's principles for guidance.

The third overarching theme in the *Statement* is that self-represented litigants are at an inherent disadvantage relative to represented parties. This claim is perhaps best rooted in criminal cases, wherein section 10(b) of the Charter provides a constitutional right to legal assistance in situations of arrest or detention (*Canadian Charter of Rights and Freedoms*, 1982). However, such a right is not extrapolated to cases outside this scope, as noted in *British Columbia (Attorney General) v. Christie* (2007, para. 17, 23-27). Pondering why legal assistance should only be awarded when 'stakes are high' invites a myriad of questions relating to constitutional rights and reform of the entire legal landscape. However, many justifications supporting a right to counsel for those in arrest or detention also apply to general self-representation cases, wherein the stakes may be immense. These justifications rest on the premise that self-represented litigants are in an inherently worse position than those with experienced and knowledgeable representation. This claim applies to the competency of self-represented litigants to represent themselves in court and how judges treat and perceive them.

The *Statement* outlines that designations of vexatiousness are to be used when the administration of justice requires it to prevent abuse of process (Canadian Judicial Council, 2006, pp. 5-6, 9). However, there is growing concern that such classifications are left to the judge's inherent discretion and are, therefore, often wrongfully applied to self-represented litigants that make innocent but erroneous mistakes. As has become too apparent,

the *Statement* fails to elaborate strictly on when such characterisations will be used and for what purpose. This leads to power imbalances in the courtroom, placing self-represented litigants at a disadvantage. More detail will be provided regarding the issue of vexatiousness later in this paper; however, this provides an overview of how the *Statement* has (or has not) dealt with it.

The Council's statement on the principles of self-representation was endorsed and affirmed in *Pintea v. Johns* (2017). Further, the Court noted that self-represented litigants could not simply be treated as equals to someone appearing with representation (Naidu and Macfarlane, 2021). While this is a positive step, numerous decisions indicate that lawyers and judges still fail to live up to the judgement in *Pintea* in self-representation cases. In *Girao v. Cunningham* (2020), the Court found that the trial judge "allowed himself to be led by trial counsel's arguments", whereas "Ms Girao, a self-represented, legally unsophisticated plaintiff who struggled with the English language, was left to her own devices" (para. 156). "Fairness required more", wrote Justice Peter Lauwers as he commented on the inadequacy of judges and lawyers that frequently fail to abide by the standards set out in the Council's *Statement* (para. 159). Indeed, he was right. Fairness requires that courts and agents of the justice system are aware of the importance of distinguishing cases wherein at least one party is self-represented. Since they face an inherent disadvantage, ensuring equitable access to justice and fair treatment requires realising that the idyllic principles do not always hold, by themselves, in practice. It is up to the lawyers and judges who command the courtroom to put these into effect.

2.2 Current State of Self-Represented Litigants

"I was standing up there and my knees were shaking and all of a sudden somewhere in the background I heard… You are fighting for your life." (Stratton, 2011, p. 89)

2.2.1 Brief Introduction

Choosing to self-represent is often not a choice at all. Many persons are forced to effectuate their 'right' to self-represent due to mistrust in lawyers, financial issues, socio-economic barriers, and structural inequities. In a system that prides itself on justice and reconciliation, many find little help or opportunities to obtain the legal services they require. Eligibility criteria for legal aid are complex and exclude those in most need of legal assistance. Unbundled legal services are insufficient to cover parts of legal matters that demand a lawyer's expertise at affordable costs. Courthouse hours, commute times, and intake and application processes pose more obstacles that self-represented litigants must hurdle. Even with the abundance of online materials and resources, there is a lack of guidance on how to utilize or apply those resources to specific cases. The digital divide between age groups and inaccess to virtual services pose significant barriers to self-represented litigants looking for ways to navigate a rigid and overly complicated legal system.

Further, there is not one type of self-represented litigant. Self-represent litigants come from all parts of society with varying ages, incomes, and education levels. There is also a field of differing motivations for why they are forced to self-represent. Many cite financial issues as the most critical factor; however, disabilities, ethnic background, mistrust in the justice system, and domestic issues play essential roles. Persons affected by these inadequacies may be further plagued by socio-economic and political factors that protract the marginalization of disadvantaged groups. These exacerbate struggles of everyday life in a system designed to

correct them. The only common thread amongst self-represented litigants is that they all seek fair and adequate access to justice, which they are currently deprived of.

There is another trend to characterize self-represented litigants as vexatious. Vexatious litigation refers to a legal action whose sole purpose is to harass or subdue the opposition. The courts frequently use this designation to label self-represented litigants as disruptors or filibusters, fundamentally devaluing and discrediting their legal position. Misuse of this characterization leads to power imbalances and inefficiencies in the court system, which fail to provide adequate compensation and redress in cases that deserve it the most. Recognizing this misjudgement's effect on the self-represented litigant, who may already struggle to cope with the courts' intricacies, vagaries, and intimidations, is imperative. The 'vexatious' characterization is reliant upon the premise that self-represented litigants are there to impede proceedings when, in reality, little thought is paid to why proceedings are hindered in the first place. It is not because of any deliberate intent to disrupt but rather because of a breeding unfamiliarity of and mistrust in the justice system that self-represented litigants are perceived as vexatious.

Self-represented litigants are on the rise, and it is of utmost importance to promote a realistic image of who self-represented litigants are and their misperceptions in the courts to rectify the systemic injustice they face. In particular, this section examines the current state of self-representation through four lenses: (1) there is "not just one type of self-represented litigant"; (2) the number of self-represented litigants is increasing; (3) they have difficulties navigating the legal system; and, (4) face issues of vexatious designations (Research and Statistics Division, 2013). Each will be discussed in turn.

2.2.2 Not Just One Type of Self-Represented Litigant

While there are certainly trends in demographics and socio-economic statuses of self-represented litigants, the view that there is one type of person that chooses to self-represent is both incorrect and detrimental. Self-represented litigants "come from all income and education levels", have differing motivations for choosing to self-represent, and can be of various ages and ethnic and national backgrounds (Research and Statistics Division, 2013). Whereas many low-income and unemployed households qualify for legal assistance, some with higher incomes are ineligible and must bear the expense of hiring a lawyer themselves – a factor that consequently means many choose the path of self-representation (Sinha, 2014; Stratton, 2011, p. 109). Those best served are wealthy individuals who can afford a private lawyer and low-income households with higher education, presumably eligible for legal assistance (McLachlin, 2008; Stratton, 2011, p. 42). Households with higher education, in general, are more often able to find, access, and take advantage of available legal services (Arshad, 2007). However, these persons account for a small proportion of those that self-represent (Farrow et al., 2012, p. 4).

Low-income self-represented litigants that lack social resources are usually affected by a plethora of additional social barriers to accessing justice, apart from financial and eligibility criteria, that include mental and cognitive disabilities, language and cultural barriers, and living in remote locations (Action Committee on Access to Justice in Civil and Family Matters, 2013; Farrow et al., 2012, p. 4). This is of particular concern as poor and vulnerable populations tend to "have high rates of intersection with civil and criminal legal problems" (Currie, 2007, p. 34; Currie, 2009; Stratton, 2011, p. 30). These legal issues can propagate and exacerbate social exclusion, income polarization, and discrimination leading to "significant costs for individuals… and society as a whole" (Semple, 2010; Stratton, 2011, p. 30; Stratton & Anderson, 2006). Therefore, the issue of self-representation affects different individuals in different ways – not just statically and exclusively one group of persons.

While financial reasons and ineligibility to acquire legal assistance are among the most common motivations for self-represented litigants, studies have shown that the decision is complex and multifaceted (Macfarlane, 2013; Sinha, 2014). Many persons that self-represent do so as they cannot find legal assistance or guidance. This disproportionately affects those living in rural areas with inaccess to services and older generations struggling to cope with the digital divide as many services move to an online model (Sullivan & Macfarlane, 2021, p. 26). Further, some choose to self-represent as they have previously had negative experiences interacting with lawyers and judges, a reason which seems to have festered into a general lack of trust in the justice system (*Dujardin v. Dujardin*, 2018; Macfarlane, 2013). This is particularly prevalent amongst Indigenous populations, LGBTQ2IAS+ persons, and those with mental and physical disabilities (Department of Justice, 2021; Royal Commission on Aboriginal Peoples, 1996; Stratton & Anderson, 2006).

Mistrust in the justice system is also characteristic of self-represented litigants involved with legal issues in their youth (Feldstein, 2016; Stratton, 2011, p. 92). These are consequences of a much broader issue in the justice system relating to children and young persons that this paper does not discuss. However, studies have found that young self-represented litigants are "routinely denied or manipulated" in the legal system, wherein a lack of support and trust leads to trauma and destabilization (Select Standing Committee on Children and Youth, 2009; Stratton, 2011, p. 92). The inability to address these issues and the seething unfamiliarity with legal structures and formalities force many young persons to reject any claim to equal and adequate justice.

While self-representation affects many persons, those living in poverty, vulnerable areas, and excluded communities tend to be disproportionately impacted as a cause of additional social barriers. Unfortunately, Indigenous populations often fit this description. A self-represented litigant identifying as Indigenous is "likely to experience overt discrimination"

(Currie, 2007; Stratton, 2011, p. 94; Stratton & Anderson, 2006). Further, research shows that Indigenous persons are disproportionately present in groups living off low incomes and facing homelessness and addiction problems (Baral et al., 2021; Patrick, 2014; Reading & Wien, 2009; Zhu et al., 2021). As these populations are most prone to experience "clustering of multiple unresolved legal and social problems" and overrepresentation in the justice system, it is surprising that a national study found that Indigenous persons are less likely to use available legal services (Stratton, 2011, p. 88).

Research shows a lack of trust in the government amongst Indigenous communities and "a perception that the justice system is not there to serve or protect Aboriginal people" (Royal Commission on Aboriginal People, 1996; Stratton, 2011, p. 94). Previous research identifies that these convictions stem from a colonial history of injustice, discrimination, and punitive experiences within the justice system – a system that self-represented Indigenous litigants are still confined by (Brant, 2020; Gunn, 2016; Loppie et al., 1986). Indigenous self-represented litigants frequently cite a lack of appropriate cultural services and training that reflect the unique needs of native communities (Action Committee on Access to Justice in Civil and Family Matters, 2013; Gutierrez et al., 2018; Stratton, 2011, p. 134). Further, inasmuch as our justice system is structured around services and organizations, this is inherently at odds with Indigenous culture. Legal aid and service agents note that Indigenous persons "build relationships slowly and with individuals, not services" (Day et al., 2012; Stratton, 2011, p. 94; Ward, 2019; Zeidler, 2011). Therefore, solutions aimed at solving issues faced by self-represented litigants need to account for the diverse body that self-represented litigants constitute. They cannot simply focus on new and improved specialized services but must accommodate the most vulnerable and desperate populations.

The answer to 'who are self-represented litigants' will likely never be definite as there is not just one type of person who chooses to self-represent. A myriad of considerations goes

into making such a decision, and an even bigger dimension of factors impacts persons leading up to choosing to self-represent. This paper will later show that self-represented litigants come from all levels of society and face additional systemic issues that intersect with many social and economic considerations.

2.2.3 Increasing Number of Self-Represented Litigants

Self-representation affects more and more persons each year (Birnbaum & Bertrand, 2013; Research and Statistics Division, 2013). Multiple studies in Canada have shown an increase in the number of persons choosing to self-represent since the start of the 21st century (Devry Smith Frank L.L.P., 2017; Farrow et al., 2012; Stratton, 2011; Sullivan & Macfarlane, 2021; Sutherland & Richards, 2017). A 2011 report found that 45% to 52% of Canadians are likely to face significant legal problems at any one time, wherein the majority "do[es] not successfully access legal assistance" (Stratton, 2011, p. 22).

Further, a 2014 survey of legal professionals, including judges, estimated an increase in the number of self-represented litigants in the past five years: ranging from 50% and 80% of all parties in civil and family law (Birnbaum & Bertrand, 2013; Boyd et al., 2014; Boyd & Bertrand, 2014). This is supported by Julie Macfarlane's (2013) findings which estimate that 40% to 57% of parties at the precept family law cases are self-represented. This number rises significantly when the case is opened, wherein family court reports reflect a number between 64% and 74% (Macfarlane, 2013, p. 33). Self-represented litigants are also becoming more and more prevalent in big cities. Statistics suggest an increase from 50% of persons in courts being self-represented from 1998 to 2003 to 80% of Toronto's litigants in 2021 (Lynch & Davis, 2021). This has undoubtedly led to a rise in demand for adequate and available legal services (Action Committee on Access to Justice in Civil and Family Matters, 2013).

In general, no arguments in the current literature suggest that the number of self-represented litigants is decreasing. However, the data that currently does exist is sparse, and this is mainly due to government agencies not tracking self-represented persons and case-related information. There needs to be a bigger collective drive by provincial and federal government organizations and courts to obtain data on self-represented litigants, their numbers, and their problems.

2.2.4 Difficulties Navigating the Legal System

Self-represented litigants face inherent issues navigating the legal system and correctly understanding and appreciating the legal structures and processes required in court proceedings. Attempts to request legal assistance further threaten to cross into legal advice, which makes lawyers and judges hesitant to provide full support (Hannaford-Agor & Mott, 2003; Greacen, 2001). This leads to power imbalances in the courtroom as self-represented litigants struggle to keep up with the knowledge and expertise judges and lawyers have spent their lives developing. Self-represented litigants must also balance their time, personal, and familial obligations with preparing and understanding the intricacies of their case. This places low-income and socially vulnerable groups who do not have the luxury of free time in precarious and desperate situations that further impact those struggling with physical and cognitive disabilities. In short, research suggests that cases involving self-represented litigants take up more court time, are more costly to all parties, and lead to unrealistic and sometimes counterproductive expectations (Greacen, 2014; Macfarlane, 2013; Research and Statistics Division, 2013).

Legal assistance agents and court staff report that their typical interactions with self-represented litigants take considerable time (Macfarlane, 2013). A 2012 white paper prepared for the Association of Canadian Court Administrators noted that most court staff interact with

two to three self-represented litigants per day – a number which used to be two to three per week (Farrow et al., 2012, p. 15). In each case, the time necessary to sit with the self-represented person is roughly two to three hours. However, since self-represented litigants are more unfamiliar with court processes and required documents, this often results in repeat contact and more time spent outside court. In court, lawyers and judges note that cases involving self-representation take up "more court time, are less likely to settle… [and] legal costs increase" (Research and Statistics Division, 2013). These trends are especially prevalent in family law cases, wherein special considerations such as custody, access, and support require more time and appearances from self-represented litigants. Despite research suggesting that self-represented litigants receive more help and are devoted more attention and time, lawyers and judges report that self-represented litigants are often worse off compared to represented parties (Macfarlane, 2013).

The unusually long hours characteristic of self-representation cases spent preparing for and performing in court lead to increased costs that affect all legal agents. It is innately at odds with the Department of Justice's vision of an efficient justice system and proliferates administrative, transactional, and bargaining costs (Cooter & Ulen, 2016). Procedural errors are also common, given a lack of knowledge of law and courtroom formalities which add to an already pressing burden experienced by court staff (Farrow et al., 2012). Failure to file pleadings, incomplete documents, and numerous scheduling adjustments further increase the costs associated with inefficient timing allocations, which trickle down onto every subsequent case – affecting both those represented and unrepresented. These costs seem to be exacerbated in cases wherein the self-represented litigant has low language comprehension and digital literacy rates. This affects "linguistic minority groups (notably Francophones and First Nations) and immigrant women" who require special assistance from the courts, a service which is not always provided (Stratton, 2013, pp. 46-47).

Multiple sources of literature also note that persons dealing with legal issues are typically under stress, sometimes in crisis, which makes rigid and formalistic legal language immensely difficult to comprehend even for highly educated persons: "most of us [the average person] don't know what to do, where to go, and who to speak with" (Access to Justice Task Force, 2009; Gander et al., 2005; Noone, 2009; Stratton, 2011, p. 6). All these considerations lead to long and complex cases wherein each day of trial is estimated to cost the average taxpayer $10,000 (Jukier, 2015, p. 220). The Action Committee on Access to Justice in Civil and Family Matters further estimated in 2012 that a 7-day trial in Canada costs, on average, between $38,000 and $125,000 (Action Committee on Access to Justice in Civil and Family Matters, 2013, p. 4). As such, it is not difficult to understand why groups with particularly low-income levels and additional social barriers struggle to gain access to justice in Canada. The difficulties of navigating the legal system and inaccess to relevant services contribute to these costs of which self-represented litigants are most vulnerable.

It is worthwhile to note that, consistent with the deficiency of information being tracked on self-represented litigants, there is no data that precisely measures the costs of self-representation to the justice system (Research and Statistics Division, 2013). While virtually all research on self-represented litigants agrees that their unfamiliarity with courtroom processes and legal structures increases costs, more data needs to be tracked to quantify this. Such data could significantly increase our understanding of the most inefficient parts of the justice system and the most cost-effective policy considerations for rectifying injustices experienced by self-represented persons.

Due to a lack of knowledge of legal proceedings, precedent, and courtroom formalities, self-represented litigants are also more likely to have unrealistic expectations regarding the outcome of their case (Macfarlane, 2013). While this may seem like an uninformative variable, expectations significantly determine the number of cases filed and exacerbate inefficiencies

and costs to the justice system (Cooter & Ulen, 2016). On the former, self-represented litigants are inherently more likely to have distorted views on the value of their legal claim, given that they have less knowledge than legal professionals. This information asymmetry leads many self-represented litigants to trial instead of settling disputes (Macfarlane, 2013; Research and Statistics Division, 2013).

Further, self-represented litigants have difficulties detaching themselves from the case to argue objectively and rationally (Farrow et al., 2012; Greacen, 2014). Represented parties are much better positioned in this regard as they are not at the helm of their own case. This leads to a proliferation of self-representation cases wherein the unrepresented party may be blind to information that legal professionals consider decisive (Greacen, 2001). While this imposes administrative and transactional costs on the justice system, self-represented litigants pay the greatest penalty for their innocent but unrealistic expectations. The asymmetry of information between parties also leads to judgements that do not reflect an efficient allocation of resources or distribution of justice. Reasons for this include self-represented litigants refusing to settle as they have an unrealistic understanding of their legal claim and then losing at trial, wherein they receive no damages or redress that perhaps should have been awarded as the most efficient and just outcome.

The courts' conventional method of dealing with such problems of unrealistic expectations has been to increase filing costs to disincentivize self-represented litigants from pursuing legal action (Cooter & Ulen, 2016). However, this policy is predicated on a financial and economic motive to accessing justice. It does little to address the root cause of the issue – that self-represented litigants suffer from a lack of knowledge and assistance in the legal system. While increasing filing costs may be less expensive from an administrative perspective, it leads to a rise in court costs and negates the potential positive benefits of promoting legal education and providing assistance – that which self-represented litigants desperately need.

While there is a lack of data on the costs of unrealistic expectations of parties in Canada, scholars concur that it is a pervasive problem mainly affecting self-represented litigants (Cerniglia, 2020; Feldstein, 2016; Greacen, 2014). A 2003 study found such results in cases involving self-represented litigants in the United States (Hannaford-Agor & Mott, 2003). In small claims court, represented parties were more likely to receive a favourable judgement after controlling for case characteristics (p. 171). Cases where both parties were unrepresented commonly resulted in dismissals for failure to prosecute. These findings were significantly normal in four of the five counties of interest they studied. As for the fifth (Ventura) county, "virtually all of the cases with self-represented parties resulted in dismissal" (p. 171). The authors conclude that these results point to two observable patterns for self-represented litigants: their cases commonly result in dismissal, and those that do not commonly result in unfavourable outcomes. Such findings seem to be perpetrated by the asymmetry of information between represented and unrepresented parties, and the unrealistic expectations of the latter (pp. 174, 178). While the study was conducted in the United States, it is consistent with the small observations available in Canada: "75 per cent of S.R.L.s lost at trial and over 85 per cent of S.R.L.s lose in motions court" (Lynch & Davis, 2021).

2.2.5 Issue of Vexatious Designations

Self-represented litigants commonly cite issues of procedural fairness as a cause of their mistrust of the justice system (Campbell & Macfarlane, 2019, p. 7). Among these issues lie designations of vexatiousness, which is a term used to denote litigants whose actions solely intend to disrupt court processes. Former Justice Minister Ron Stevens described vexatious litigants as persons who persistently bring meaningless legal claims, fail to pay costs, abuse courtroom decorum, and engage in opinionated and narcissistic behaviour (Lieb, 2010, p. 74). The term is often incorrectly used to describe self-represented litigants whose innocent

mistakes and misunderstandings are conflated with attempts to filibuster and abuse the court's jurisdiction (Campbell & Macfarlane, 2019, p. 8). Such designations lead to power imbalances in the courtroom that directly affect self-represented litigants' experiences dealing with the justice system (Shushani & Macfarlane, 2018, p. 2). The lack of judicial oversight and inconsistencies among judges' characterizations of vexatious litigants is particularly concerning, which propagates the marginalization of those unrepresented.

Studies report that formal vexatiousness is being used at the judge's 'inherent discretion', which is a legal term used to denote a judge's right to authority and control over the courtroom without needing to reference a rule or statutory provision (Campbell & Macfarlane, 2019, p. 8; *Hok v. Alberta*, 2016, paras. 14-25). The lack of standardized practices and guidelines for imposing such a designation leads to efforts to characterize extreme behaviours by self-represented litigants as formally vexatious, wherein they may be products of unintended errors. Even in some cases where the judge does not formally term behaviour vexatiousness, there have been instances of 'vexatious lite' designations (Macfarlane, 2021).

Importantly, vexatiousness is becoming less concerned with issues of procedural fairness or intentionality on the self-represented litigant's part. Instead, the term enables focusing on mistakes and misunderstandings that self-represented litigants are disproportionately more inclined to make. This also seems to cut across gendered lines, wherein females are described with language that invokes a power imbalance in the courtroom and displaces them from a position of equality (Shushani & Macfarlane, 2018). Terms such as 'schemers', 'dramatic', and 'attention-seeking' all portray the female self-represented litigant in a negative way that undoubtedly harms (though has nothing to do with) the merits of the case (*See C.L.M. v. M.J.S.*, 2017, para. 9; *Ottewell v. Ottewell*, 2013, para. 7; *S.L.M.D. v. A.V.D.*, 2017, para. 44). Males are also sometimes described as combative and abusive which protracts

the marginalization of self-represented litigants in the courtroom who are only looking for equal footing (Campbell & Macfarlane, 2019, p. 10).

There needs to be more significant consideration on the lawyers' and judges' part regarding the specific backgrounds and socio-economic positions of self-represented litigants. Notably, courts should not be able to systemically punish self-represented litigants who make unintentional errors or innocent mistakes, especially not by imposing costs to an already expensive process which some courts have done (*De Silva v. Fraser Health Authority*, 2013, paras. 69-70). Instead, the inherent jurisdiction of the judge in self-represented cases needs to be scrutinized and subject to oversight. The Court of Appeal in Alberta, for example, restricted the use of inherent jurisdiction in *Jonsson v Lymer* (2020, para. 44). The Court detailed that designations of formal vexatiousness should only be reserved for extreme situations in self-represented litigants cases, wherein the behaviour represents an unmistakable pattern of abuse to the courtroom process. *Jonsson* noted many other reasonings that concur with the sentiment of this paper, such as special consideration for self-represented litigants and more awareness of judges' roles. However, on vexatiousness, it makes clear that the term presents an issue to self-represented litigants looking for access to justice, and more attention needs to be paid to how to address it.

3. Analysing Demographics and Experiences of Self-Represented Litigants:

National Self-Represented Litigants Project Data

This paper now examines the experiences of self-represented litigants through an original analysis using data made available by the National Self-Represented Litigants Project (NSRLP). The data provides space to investigate self-represented litigants' socio-economic and demographic characteristics. The exploratory data analysis presents these statistics and concludes with observable trends to quantify who the self-represented litigant respondents are. Whereas the first portion of this paper examines existing research and literature, the data here aggregates results from an ongoing survey by the NSRLP to extrapolate sample time-variant and idiosyncratic effects on the general state of self-representation in Canada. The following predictive and inferential analyses quantify self-represented litigants' experiences through machine learning and regression models to determine whether experiences are correlated with demographic and socio-economic trends. This is done through sentiment scoring, which provides an exciting avenue for future research. Overall, this paper strongly promotes incorporating quantitative analyses into evaluations of the self-representation problem, which is imperative to pursue evidence- and data-based policies.

As a brief note, this section has reduced profound statistical commentaries to a minimum to make them more digestible for a general audience. However, because the methodology relies heavily on statistical theories, and in best practice, these considerations are still introduced. To make the statistics more understandable, I use first-person and narrative writing to explain the intricacies behind the methodology and models. I aim for you, the reader, to take away trends in self-represented litigants' experiences and, specifically, the factors that interrelate with or provoke a bad experience. Hopefully, you will also better understand how new and exciting statistical models and machine learning techniques can quantify qualitative data and be used to examine the self-representation issue in future research.

3.1 Data and Exploratory Analysis Methodology

The data in this analysis is taken from the Intake Form of the National Self-Represented Litigants Project. Specifically, 768 responses from self-reported self-represented litigants were collected from 2016 to January 2023, when this analysis was conducted. The Intake Form asks the respondent a series of multiple choice and long answer questions that, given the respondent's consent, collects socio-demographic information such as income, age, gender, ethnicity, and education levels. Further, the Form records short- and long-answer responses that detail the experiences of self-represented litigants and associated variables. Importantly, all responses used in this analysis are anonymized to protect the respondent's privacy.

In order to correct inconsistencies in form responses and process the data for computational analysis, the data is manually cleaned and encoded using tidyverse and ggplot2 packages in R. 'Cleaning' data refers to the process of preparing data for further analysis. Notably, programming languages and statistical software cannot infer the correct response type from spelling errors or textual responses. Therefore, it is essential to first translate these answers into data understandable for computers and imputable in statistical functions. To do this, I remove a subset of irrelevant variables (questions that were asked of the respondents) from the complete dataset that provide no valuable addition to our research. Such variables include respondent contact information and location, which mainly serves to anonymize the data. Second, upon examination of the data, I notice some responses labelled 'TEST' or 'testing'. These observations are subsequently removed as they are presumably inputs from the form administrators and provide no information on the experiences or characteristics of self-represented litigants. Standardizing form responses is the third and most computationally heavy step in the data pre-processing component. Standardizing responses includes correcting spelling errors so that respondents whose case was tried in 'Ontario', 'ON', and 'Ontaraio', for example, converge. For the most part, simple response errors are easy to correct as they indicate

a particular category, such as a specific province or territory in the example above. However, ambiguous responses, such as 'j', 'IA', 'dsd', and 'g', were labelled undefined. This process is repeated iteratively for all observations and variables, including correcting classifications of gender (male, female, non-binary), income levels, education levels, and long-answer responses. Questions are also re-encoded to be more feasible for data analysis so that the variable "Do.you.identify.as.a.person.with.a.disability." becomes "disability" with a simple yes or no binary output. This is similarly done for all binary variables so that yes is indicated by the number 1 and no by 0. Note that 'binary' simply means that the response can only have one of two answers: typically, yes or no. This is important because, while appearing extremely sophisticated, computers are actually quite simple, running on mass amounts of data characterized by 1s and 0s (although interpreting those is not quite as simple!). Because of this, encoding yes or no responses as either 1s or 0s allows the computer to understand better what we are inputting. In a human's mind, for example, this process is like changing an overly complicated sentence: "I am disinclined to acquiesce to your request" to a much simpler "No." Finally, before moving on to the exploratory data analysis, the cleaned and processed datasets are exported to CSV format for use in regression and classification exercises later on.

The exploratory data analysis (EDA) provides essential information about our population of interest (self-represented litigants), which should be examined before conducting more sophisticated methods to evaluate assumptions underlying statistical models. Conducting an EDA is imperative to understand who the respondents are and whether our data can be extrapolated to the general population of self-represented litigants. It also functions as a replacement for a summary statistics table which is considered the norm, however unfeasible to generate in our case due to the categorical nature of responses.

3.2 Methodology for Sentiment Analysis

To better understand self-represented litigants' experiences and infuse a data-oriented approach into a conventionally qualitative area, this report executes a sentiment analysis using classification machine learning models to correlate long-answer responses on experiences with quantifiable impacts. Woah, what does that mean? Essentially, computers are notoriously horrible at understanding text. While we rely on natural languages to interact with one another, computers are generally incapable of understanding a worded language not characterized by 1s or 0s, as noted above. Those who observed the rise of Chat-GPT might be wondering why that application can understand natural language text inputs relatively well. Well, the methodology below might help explain, but, in short, Chat-GPT is also a machine learning model similar to the ones used in this analysis. To rephrase the opening statement in simpler terms, this part of the analysis aims to teach the computer how to read long-answer responses on self-represented experiences and classify them as either good, bad, or neutral.

The first step in teaching a computer how to 'read' is to break down the text data into numbers. While this may sound counterintuitive, the computer has its own dictionary to look up which numbers are associated with which words: almost as if it is translating the text into its own language (which only computers can understand). From there, statistical models infer results based on the occurrence, dependence, and structure of a set of words, i.e., sentences. Fortunately, a large variety of packages (think 'dictionaries') exist that a computer can use depending on the types of texts we want it to read. This analysis uses the NLTK and VADER packages instead of spaCy, Gensim, Pattern, TextBlob, Scikit-learn, and Pytorch, to name a few. The justification for this is that NLTK and VADER are oriented around short and non-contextual Twitter-like texts, similar to the long-answer responses collected by the Intake Form. Whereas other packages, such as spaCy or Gensim, are excellent, they are more industrial and require more extensive texts to extract context and meaning. Therefore, NLTK and VADER

provide a solid foundation for this analysis, but not every analysis, which we can use to teach the computer to essentially read the responses in its own language.

Hopefully, this clarifies where the process begins and *how* computers can read texts. The second step involves figuring out *what* computers can read. As mentioned earlier, the computer looks at a set of words' occurrence, dependence, and structure – what we conventionally refer to as sentences. However, because the computer needs to translate every word into its own language, some things may get lost in translation. Therefore, we want to give the computer the most specific and relevant parts of a sentence to minimize the risk of losing the most crucial part. To do this, statisticians use a method called 'pre-processing' that essentially converts our original sentence to its simplest form without removing the important words. This process involves text normalization, such as 'tokenizing' and 'lemmatizing' sentences. The former refers to splitting the sentence up into individual words so that "The dog ate my homework!" becomes ["The", "dog", "ate", "my", "homework", "!"]. The latter, lemmatizing, conjugates words into their root or simplest meaning: ["The", "dog", "eat", "I", "homework", "!"].

While this sentence may not make much sense to us, the computer understands it perfectly. However, remember the importance of including the most relevant parts of a sentence. Words like "The" and "I" and punctuation do not add much value to a computer's train of thought. Therefore, statisticians remove these words referred to as 'noise', just like an artist or producer would remove or reduce white noise in their audio recording. Therefore, our end-product is ["dog", "eat", "homework"]. Hopefully, it is clear why this process simplifies sentences while keeping the critical parts. This process clarifies *what* we want the computer to read, and it can now easily look up what "dog", "eat", and "homework" is in its own language while maintaining the meaning of the sentence.

The final step is conducting the actual analysis. This section is interested in whether

respondents generally describe their experience of self-representing as either good, bad, or neutral. This investigation is what statisticians call performing a 'sentiment analysis.' Incredibly, the computer also has a whole other dictionary for the sentiments of each word – which tracks intuitively with words we consider to be either good, bad, or neutral. Let us pick another example. The pre-processing step explained above would convert the sentence "This doesn't feel very fair to me" to ["not", "fair"]. Remember, we remove any words that do not add significant meaning. You, the reader, would presumably classify the previous sentence as generally negative. However, the computer reads "not" as negative but "fair" as positive. Aggregating these would mean that the sentence is neutral? This doesn't seem right.

Interestingly, while computers interpret each word individually, statisticians use a method called N-gramming to map the dependence and structure of a set of words to specific meanings. This allows the computer to interpret multi-word statements together. In this case, we have "not fair", a set of 2 words. Therefore, we can get the computer to fully understand the sentence using a *bi*gram approach. The computer now looks at ["not", "fair, "not_fair"] to evaluate the sentence. By reading the two words together, the computer can determine that "not_fair" actually means unfair, which is negative. A few moments later, the computer informs us that that set of words actually indicates a negative sentiment. In short, that is how a computer determines that the sentence we are interested in ("This doesn't feel very fair to me") is negative. By applying this exact process to the long-answer responses, we can obtain a sentiment score for each respondent on a continuous scale from -1 to 1, where negative numbers mean a negative sentiment, positive numbers mean positive, and between -0.05 and 0.05 means neutral.

3.3 Methodology of Machine Learning Classification Models

Now that you understand the basics of textual and sentiment analysis, we can begin to explore what is possible to infer from the results. Specifically, since we have classified our long-answer responses as either negative, neutral, or positive, it would be interesting to see if the computer can detect a relationship between our sentiment scores and other short-answer and binary variables, such as income, age, education, and ethnicity. For example, would a hypothetical Black, 40-year-old, university-educated self-represented litigant with an income of $50,000 be more likely to describe their experience as negative, neutral, or good? This process is called 'classification'; unsurprisingly, given appropriate data, computers are generally quite good at this. Here is where the 'machine learning' component starts to take over.

Machine learning essentially means that the computer can detect patterns in data independently without us having to spoon-feed its instructions – the machine learns! However, sometimes computers do not learn that well, similar to how one would perform poorly on a biology test if one studied physics. Therefore, the computer is given a subset of the data (usually 80%) to learn from. We then evaluate whether it learnt everything correctly by examining its performance on the remaining 20% of our data. However, the trick is that the 'test' dataset does not include the answers – after all, we want to see if the computer can predict them, just like a person writing an exam. The computer studies 80% of the questions and then tries to figure out the remaining 20% independently. By determining how correctly a computer classifies the results, we can uncover patterns in the data that may be invisible to the naked eye. This is, in brief, what machine learning is and how this analysis will use it.

Just like how one person may perform better on tests relative to another person, some machine learning models perform better than others. Therefore, it is important that we apply multiple different machine learning models in order to find the ones that work best. In this analysis, we train four different machine learning models on 80% of randomly selected

responses and compare how well each model correctly classifies the remaining 20% of randomly selected responses as either good or bad. Remember that computers only understand 1s and 0s. Therefore, we can only classify something as 1 (good) or 0 (bad). The four models explored are a decision tree classifier, a random forest classifier, a logistic regression classifier, and a KNN classifier.

Explaining the algorithms behind each model would, unfortunately, be out of the scope of this paper and, frankly, take up tens of pages with heavy-duty statistical terminology and methods. However, in short, a decision tree classifier conducts thousands of tests to see which values of the predictor variables (income, gender, etc.) are related to each outcome: good or bad sentiments. It then predicts the outcome using the threshold values – exactly like following a flowchart where the endpoints are good or bad. Essentially, it *decides* which *branches* to go down based on predictor values – hence, *decision tree*. A simple example of a decision tree has been included in Appendix A for reference. A random forest classifier is similar, except that it uses thousands of random samples of the training data and predictors to evaluate each branch. Intuitively, it is like applying thousands of decision *trees* to thousands of *random* variations of the data – hence, *random forest*. Statistically speaking, this reduces the variance of each decision tree, making the result more accurate.

Logistic regression and KNN classifiers are more complicated. However, to keep it brief, logistic regression classifiers create a straight line that divides the data into two zones depending on which predictor values are associated with each outcome: in our case, good or bad. A KNN classifier also creates a line; however, it is not straight. This model looks at 'neighbours', which are responses that share similar values. Specifically, we can tell our K-nearest neighbour (KNN) model to classify a test point depending on what the outcomes are for the nearest K = 2,3,4…N neighbours. Therefore, it does not create a straight line but rather clusters' neighbourhoods' that are either related to good or bad outcomes. Again, a diagram

depicting the difference between the logistic regression and KNN classifier decision boundaries is included in Appendix B. It is important to note that each model has different applications, and thus we can never rule one out before comparing how it performs relative to other models. I hope you now understand the basics behind machine learning and some different ways a computer can learn (through different models).

3.4 Methodology for Regression Analysis

Regression models differ from classification models because they do not try to predict anything. Instead, we leave the predicting for ourselves. Regression models infer the strength and significance of relationships and patterns in data. Therefore, we can quantify to what extent one variable affects another. For example, returning to the sentiment scores, we are no longer interested in simply good, bad, or neutral outcomes. We want to quantify *how* good or bad a response is. Note that this is not a predictive exercise, like the classification process above. Instead, it is inferential and focused on extracting causal effects from the data.

Specifically, we want to determine how good or bad a response is given age, ethnicity, income, etc. Again, to steer clear from statistically heavy proofs, I will opt for a more intuitive albeit rudimentary description of regression analysis in our case. The dependent variable, which would be the continuous sentiment score from -1 to 1, is the variable that is *dependent* on others. As such, we want to examine *how* it changes depending on the values of our independent variables – those that do not change. These are age, income, and education, for example. Indeed, regression analysis allows examining how these independent variables change our dependent variable. This extracts patterns and relationships in data, from which we can infer that income affects sentiment this way, and education affects sentiment that way. In some cases, there may not be any significant relationships, and we instead find out that income does not affect sentiment, for example. However, that is for the results section to reveal.

It is important to consider threats to the internal and external validity of the regression models. Internal validity refers to how well the model performs in our sample, whereas external validity implies how well the model performs out-of-sample, i.e., in the real world. It is almost impossible to reduce these errors entirely as there will always be some level of randomness in data, which statisticians call 'noise' (from before), or innovations or residuals from one data point to the next. However, whereas some errors are irreducible, some are reducible. These depend on how the regression models are specified. To account for these considerations, the regression models in this analysis utilize 'control' variables that control for factors outside our analysis.

For example, an Indigenous person might report significantly worse experiences self-representing than a visibly Caucasian person. It would then be easy to conclude that this effect size is a cause of race, as that is the variable under investigation. However, note that Indigenous persons are often overrepresented in lower-income brackets relative to Caucasian persons. Therefore, we cannot conclude that the effect size is *entirely* a factor of race, as correlations with lower incomes might account for part of the negative experience. Income is factored into the model to rectify this 'omitted variable bias' issue. By doing so, it is possible to see precisely how much of the negative experience is a product of racial factors versus income-related factors. This process substantially improves the inferential and causal powers of the model.

4. Results on EDA and Analysis of Self-Represented Litigants' Experiences

4.1 Exploratory Data Analysis

Figure 1. Case Location.
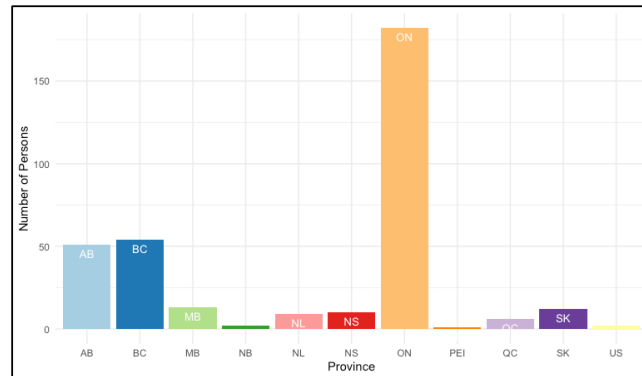
To start, as shown above in Figure 1, most respondents cited Ontario as the location where their case was held. This is intuitive as the National Self-Represented Litigants Project is primarily based in Ontario. Therefore, it is likely that most respondents who got in touch with the NSRLP were also from this area. It could also prove that the NSRLP has become the most prominent in Ontario through various efforts and outreach programs centred in this jurisdiction. However, it may benefit future NSRLP initiatives to either (1) expand awareness of the organisation in other provinces and territories, since the self-represented litigant issue is particularly stark in Manitoba and Alberta, or (2) concentrate efforts in Ontario to support their involved community and user-base.

Figure 2. Court Type.
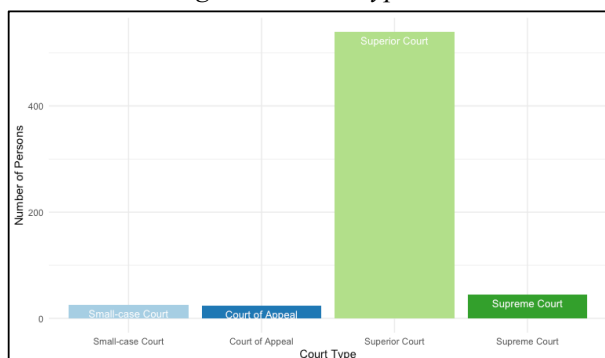
Figure 3. Case Type.
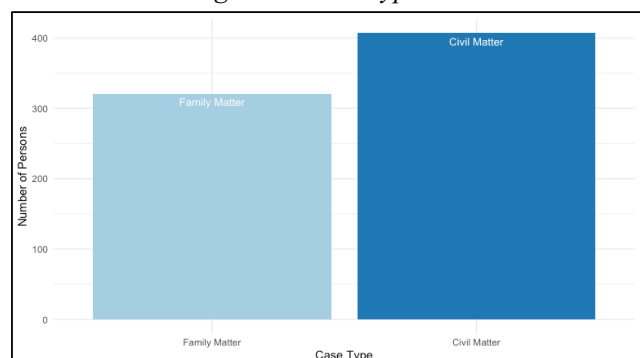
Figure 2 above highlights that most respondents interacted at the Superior Court level across provinces. However, a few respondents noted being involved in cases at the Supreme Court level. Further, Figure 3 shows that the observations were somewhat evenly distributed across family law and civil matters. This is particularly interesting as it means our results are representative of these two legal areas – and not heavily skewed towards one or the other. This allows us to extrapolate our results to a much broader population, which is critical in fully understanding the self-represented litigant problem.
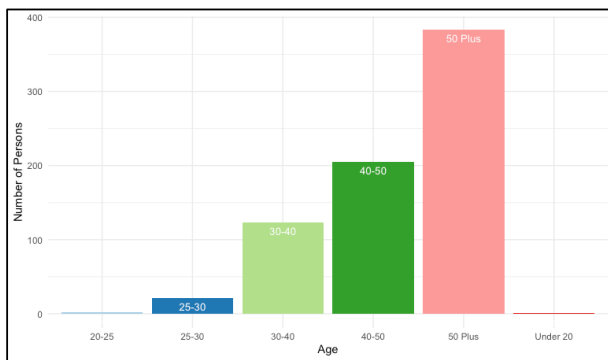
Figure 4. Age.	Figure 5. Support Friend.
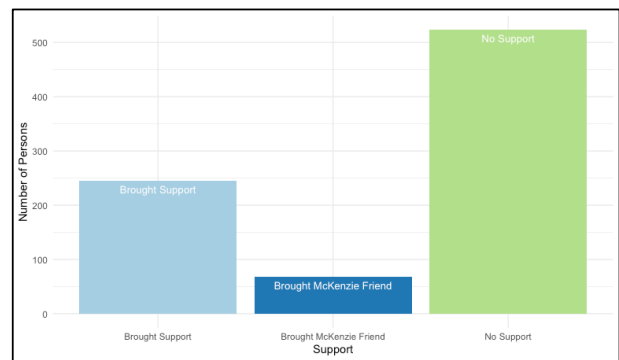
From Figure 4, it is evident that most respondents in the sample population are above 50 years old. This is interesting as the literature suggests that self-representation affects persons across all age levels. A possible reason could be that persons above 50 years are more inclined to participate in the survey because they may have more free time or feel more strongly about their experiences. It is plausible that this also correlates with general trends that older generations are more adversely affected by a lack of knowledge on accessing online resources and legal assistance through digital divides. However, further research needs to be done to corroborate this claim. Figure 5 shows that the vast majority of self-represented litigants did not bring a support person to court. While, again, the data does not provide any significant justifications for this, there may be a general unawareness surrounding the possibility of bringing a support person along to court. The data suggests this could be a potential avenue for

the NSRLP to explore further. The organization could increase the output of resources available that raise awareness of the significance and explain the positive effects of bringing a support person to assist in court.

Figure 6. Education Levels.
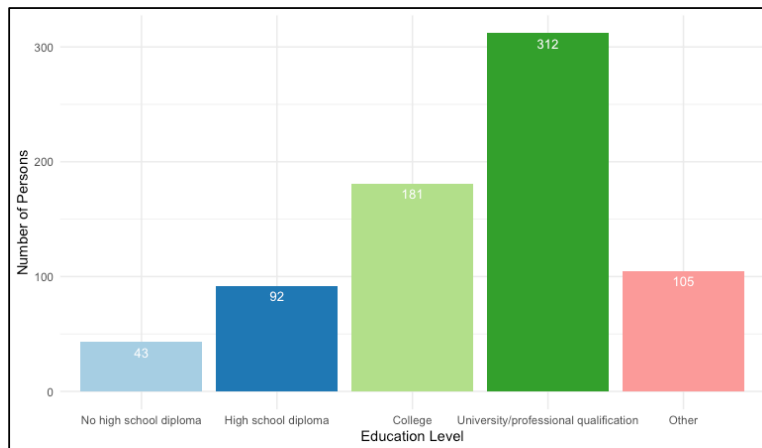
Interestingly, Figure 6 below shows that most respondents have at least a university education or equivalent professional qualification. While this seemingly refutes the broad characterization that self-represented litigants are uneducated and unsophisticated, the validity of this claim could be questioned because of potential correlations with age. Since our sample is primarily made up of persons above 50 years old, they have likely had more time and opportunity to pursue higher education. Regardless, while reasons are not conclusive by examining the exploratory data alone, the high education levels provide interesting correlations with negative experiences of self-representation, as will be discussed further below.

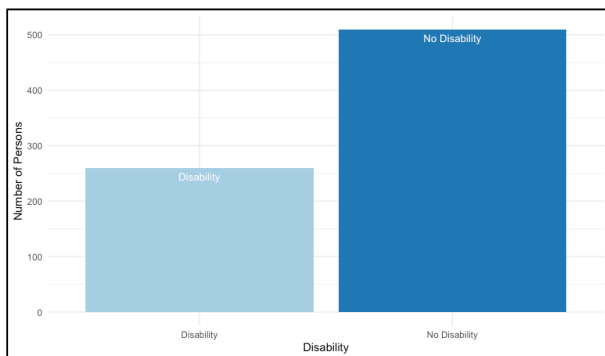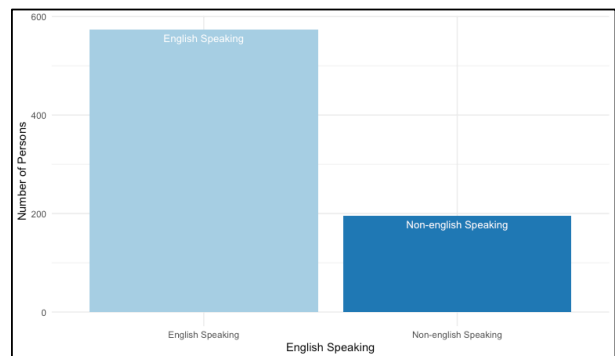*Figure 7. Disabilities.*   *Figure 8. English Speakers.*

Figure 7 highlights that most respondents in the sample do not suffer from mental or physical disabilities. However, there is still a large number of respondents that do, equating to roughly 1/3 of the sample population. This proportion of data is significant as it enables exploring the intersection of self-representation and disabilities. Figure 8 details that a relatively significant minority of the sample is non-English speaking. However, similar to the limitations of Figure 7, no data quantifies to what extent these cause barriers to fairness and what type of disability/non-English language is most present in our sample.
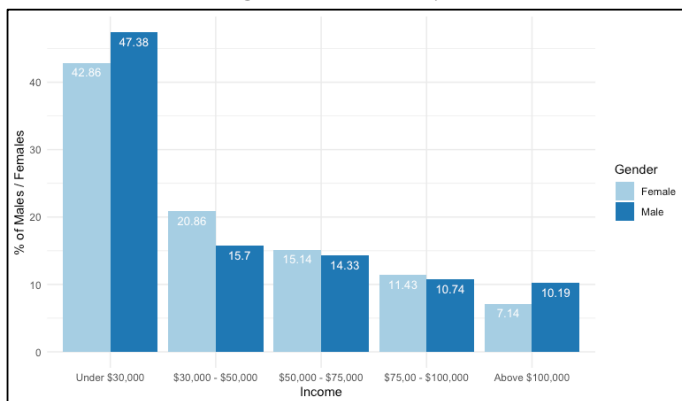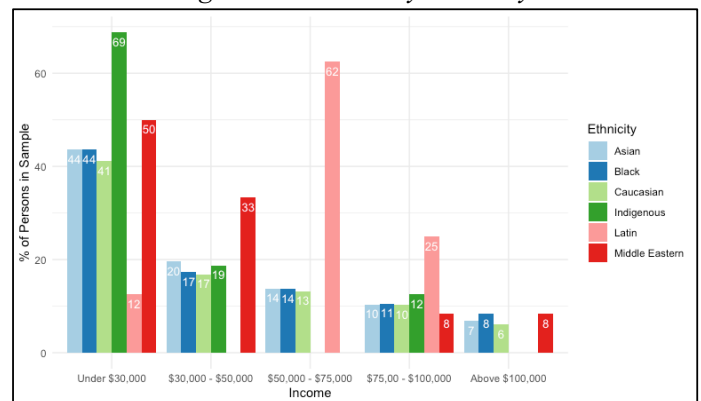

Figure 9. Income by Gender.

Figure 10. Income by Ethnicity.

Figures 9 and 10 contain income information stratified along gender and ethnic lines. First, it is interesting to note from Figure 9 that we have an even distribution of persons who identify as female or male in our sample. Only three respondents identified as non-binary, which is why their inclusion in the plot would be insignificant – however, their incomes were distributed around the medium range. Notably, Figure 9 shows that most respondents have an income below $30,000. This is significant given the expensive costs of pursuing or defending against legal action. The number of respondents falls sharply in subsequent income brackets, indicating that lower-income litigants may be more adversely affected by negative experiences of self-representing and therefore participate in the survey – which is consistent with the literature. Figure 10 shows that the data is also representative of multiple different ethnicities. However,

examining correlations by income indicates that Asian, Black, Middle Eastern, and Indigenous persons are grossly overrepresented in lower-income groups. Part of this could be the proportion of respondents belonging to a particular ethnic group which could skew our results. For example, our data contains information on six respondents that identify as Indigenous. If we were to collect more data from Indigenous persons, the result could be more evenly distributed across income groups. This could also explain the spike in Latin respondents in the $50,000 to $75,000 bracket. Nevertheless, it is consistent with literature and social science evidence that Indigenous persons are overrepresented in lower-income groups – as the EDA seemingly corroborates.

Finally, quantifying the experiences of self-represented litigants is particularly hard, given the subjective nature of the topic. However, Figures 11 and 12 below show the respondents' self-reported satisfaction with using unbundled legal services or representation.

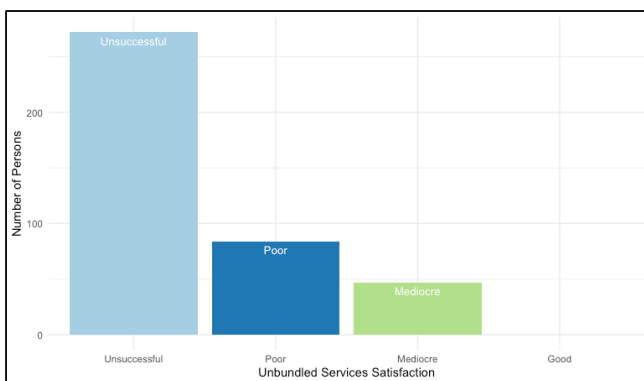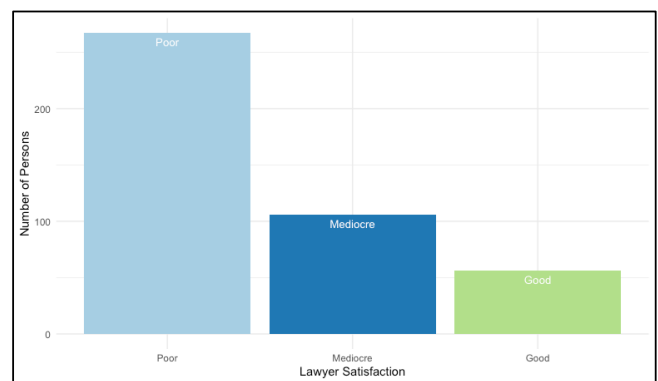| *Figure 11. Satisfaction with Unbundled Services.*	*Figure 12. Satisfaction with Legal Representation.*

It is particularly interesting to note that no respondents indicated having a good experience using unbundled legal services – the vast majority noted either being unsuccessful in attaining the service or having a poor experience. This infers that the eligibility criteria for unbundled services may be too restrictive, the service may not be offered by all legal assistance agencies, or self-represented litigants are unaware that the service exists. Figure 12 shows that most respondents represented for part of their case, either through private means or legal aid, had a

poor experience. Less than a quarter of respondents noted having a mediocre experience, whereas less than 10% cited having a good experience. This tracks with evidence outlined in Part 2 of this report, that interactions with lawyers may often lead to negative experiences due to inaccessible, impatient, non-specialized, or culturally inappropriate services.

Overall, the exploratory data analysis results indicate that our sample population seems more or less representative of the actual self-represented litigant population. Before I move on to the results of our classification and regression exercises, I want to note that our exploratory data analysis and sentiment analysis are not what we are interested in. As such, they will not be discussed here. This is because they are not inferential but rather a means through which we can conduct our primary analysis. The exploratory data analysis covered above enables an examination of what our data is trying to represent. However, it does not map relationships or the significance of certain variables.

It is nonetheless invaluable, as understanding our data is the first step in conducting any statistical analysis. On the other hand, the sentiment analysis derives the sentiment scores for the long-answer portions of the Intake Form that detail self-represented litigants' experiences. These sentiment scores are then used in the classification and regression exercises to extract patterns and infer conclusions. Therefore, while necessary on their own, they add a new dimension to the analysis that allows us to quantify how demographic and socio-economic considerations impact self–representing experiences – making them invaluable in combination with inferential and prediction-based methods.

4.2 Classification Results

Table 1. Classification Results

Model	Accuracy Score
Decision Tree	0.567
Random Forest	0.550
Logistic Regression	0.621
KNN	0.510
Random Chance	0.500

The results of the classification exercise in Table 1 show that the logistic regression classifier outperforms the other machine learning models by far. Note that the last table entry indicates 'Random Chance', which provides an intuitive understanding of what the accuracy score represents. A model that randomly classifies each observation as good or bad would have an accuracy score of 0.5. This means that, on average, a model randomly assigning outcomes with a 50/50 chance will correctly classify 50% of the observations. While an accuracy score of 0.621 may not seem like a lot in this context, it is incredibly significant. In each case, our machine learning models outperformed random chance, implying an underlying pattern or relationship in our data between self-represented litigants' experiences and their demographic and socio-economic characteristics. These models never aimed to build a predictive algorithm that could correctly classify 100% of observations. Instead, this analysis is interested in whether the machine learning methods can extract a relationship between variables – which they can.

The figure on the next page depicts the model diagnostics and performances relative to random chance visually.

Figure 13. Classification ROC Curves.
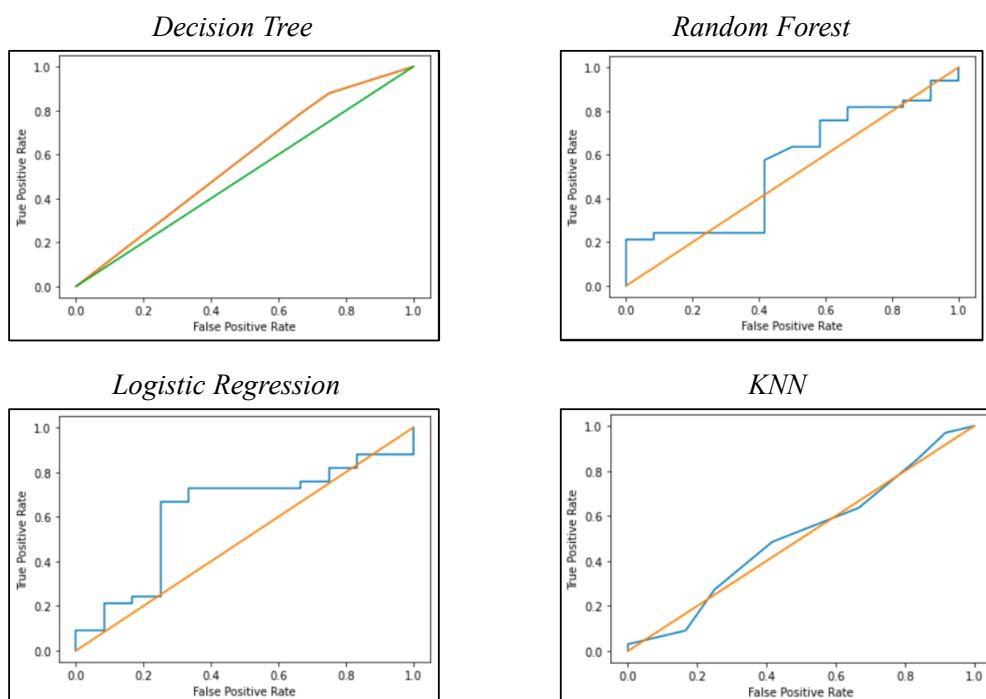
The performance of random chance is marked by the straight line in each plot. While the curvatures of the model lines matter less, statisticians tend to look for the model that reaches a 0.7 True Positive Rate first, which would be the logistic regression model in our case. Notably, the decision tree and KNN classifiers perform poorly, followed closely by the random forest model. Interestingly, the logistic regression model spikes at a False Positive Rate of ~0.25, indicating that the model correctly classifies many more observations than incorrectly classifying. The confusion matrices for the logistic regression and KNN classifiers are included in Appendix C for those of you who particularly interested to see where the models got *confused*. However, since we are not interested in the predictive power of these models, a closer examination of the matrices is out of the scope of this analysis. Nevertheless, they corroborate the above results, proving that the logistic regression model performs best. Next, we examine what we can infer from how it performs.

Having selected the best-performing model, we extract the most significant predictors that influence whether an observation is classified as good or bad. This determines which variables, such as income, education, or ethnicity, are most important to predict whether a hypothetical self-represent litigant has positive or negative experiences. The results are shown below in Table 2.

Table 2. Most Significant Predictors

Predictor	Effect on Sentiment Score
Income $100,000 +	0.853
Age 30 – 40 Years	0.397
Legal Aid Lawyer	0.316
Age 40 – 50 Years	0.229
University Education	0.122
Pro Bono Lawyer	0.079
Opposing Party has Representation	-0.0009
Family Law	-0.0968
Income $50k - $75k	-0.309
High School Education	-0.340
Income $30k - $50k	-0.392
Ethnicity Asian	-0.437
Unbundled Legal Services	-0.456
Non-English Speaker	-0.680

The results from the table above are the 14 most important predictors in determining whether a self-represented litigant had a positive or negative experience, ordered by their effect on sentiment score. Unsurprisingly, the predictor with the most considerable positive effect on sentiment score is having an income above $100,000. Notably, the effect size is 0.853, which

is substantial relative to the scoring range [-1, 1], accounting for approximately 42.6% of the space within that range. This implies that a litigant who suffers the worst experience self-representing and thus is assigned a score of -1 would instead have a mildly negative experience (score of -0.147) if they suddenly gained an income above $100,000. This finding supports the assertion that costs matter substantially in pursuing legal action. Ample disposable income safeguards against negative experiences as it undoubtedly alleviates financial and mental burdens.

It is also notable that older age seems to impact sentiment scores positively. Possible reasons for this could include the knowledge or experience that comes with age and the lesser risk of being perceived as vexatious. Notably, age between 30 and 40 years had a more substantial positive effect on sentiment scores than age between 40 and 50. This suggests that a diminishing effect may correlate with ages above 50 years, such as lacking energy, relevant services, or support networks that can alleviate stress and workload associated with the administrative and legal processes. Further, seeking assistance from legal aid lawyers positively affected sentiment scores, meaning that being eligible for such services could significantly improve the experiences of self-represented litigants. University education is also positively correlated, which seems intuitive as this enables a better understanding of courtroom processes, analytical and critical thinking, and awareness and comprehension of legal resources.

Obtaining a Pro Bono Lawyer positively predicted the good experiences of self-represented litigants; however, this effect was marginal. Similarly is the effect of the other party being represented small, however, impacted sentiment scores negatively. This could imply that the cause of the self-represented litigant problem lies inherently with issues afflicting self-represented litigants – not external factors. Whether the opposing party is represented or not seems to have a small effect, suggesting that the true variables affecting experience concern

the self-represented litigant themselves. Further, family law cases seemed to be negatively correlated with sentiment scores. While this effect was also relatively small, it could highlight the stress associated with family law cases that deal directly with personal and familial relations – often taking the most significant toll on self-represented litigants.

Unsurprisingly, having a lower income seems to affect sentiment scores negatively. This supports that income and wealth remain important for self-represented litigants who are encumbered to navigate a costly and inefficient legal system. Lacking the financial resources to do so correctly will undoubtedly lead to a worse experience. Notably, this list does not represent the lowest income group (below $30,000). A plausible reason for this could be that these populations are eligible for legal aid and assistance, which results in a more positive and generally better experience of self-representing. Having only a high school education seems to predict a negative experience, in general. This highlights the importance of awareness and understanding of legal material, which often involves analytical and critical thinking developed in further education. The fact that Asian ethnicity has a negative correlation with sentiment scores could be because of cultural differences that amount to negative experiences. However, it is also plausible that this is correlated with omitted variables, such as income, since our EDA shows that lower incomes mainly characterize Asian respondents.

Further, using unbundled legal services resulted in a prediction of bad experiences for self-represented litigants. While this is puzzling, a possible reason could be that self-represented litigants have unrealistic expectations of what unbundled legal services encapsulate. This asymmetry could manifest in a worse experience as self-represented litigants believe they are entitled to more or fail to get help in the areas they need help in the most. Finally, non-English speakers are predicted to have the worst sentiment scores. This is consistent with what one would imagine as legal language and processes are complicated and require an advanced level of English or French to understand in most cases.

4.3 Regression Results

The statistically significant results from the regression analysis are included in Table 3, shown on the next page. Notably, many variables are considered to have a relationship with sentiment score. In contrast to the classification exercise, this analysis allows us to infer patterns in the data that are representative of our sample population, i.e., the self-represented litigants that filled out the Intake Form.

First, female respondents had, on average, a lower sentiment score relative to males by -0.163. This was statistically significant at a 1% level, meaning that if we re-sample our data from another population of self-represented litigants, we would expect a different result in less than 1% of cases. Ergo, it is highly significant. The effect size is relatively large, emphasizing that female self-represented litigants are more likely to associate experiences interacting with the legal system with a negative sentiment relative to males. This highlights how the self-represented problem is particularly stark for women, who often face power imbalances in the courtroom.

Further, having an income below $30,000 results in, on average, a sentiment score of -0.112, less than persons in higher income brackets. This finding provides quantitative evidence for the claim that income is a significant variable in determining self-represented experiences. As the result was significant at a 10% level, we reject the null hypothesis that income does not affect sentiment score. Those under 30 also had a lower sentiment score of -0.645 relative to older self-represented litigants on average. This adds to our hypothesis formulated in the preceding section that age is parabolically related to sentiment, so lower age levels lead to worse experiences. However, at a certain point, this trend reverses, and becoming older also leads to worse experiences. Our classification analysis suggests that this threshold is around the 30-50 age range; however, further research should be done to corroborate this finding. It is also notable that the interaction term between income under $30,000 and age below 30 was

Table 3: Regression Results (No Controls)

	Effects on Sentiment Score
Female	-0.163***
	(0.0563)
Income under $30k	-0.112*
	(0.0571)
Age <30 years	-0.645***
	(0.0411)
Income under $30k X Age <30 years	0.655**
	(0.288)
No Education	-0.297***
	(0.0901)
Disabled	-0.0972*
	(0.0582)
Indigenous	0.433***
	(0.0677)
Asian	0.199*
	(0.103)
Legal Aid Lawyer	-0.191**
	(0.0902)
Online Service	0.140**
	(0.0622)
Observations	448 448 448 448 448 448 448 448 448 448

Standard errors in parentheses

The regression results above show the significant variables correlated with sentiment score. Notably, the number of observations dropped from 768 to 448 as not every respondent provided a long-answer response to comment on their experience self-representing.

* $p < .10$, ** $p < .05$, *** $p < .01$

statistically significant. The interaction term represents self-represented litigants who belong to both groups. The result indicates that those who do have a higher sentiment score of 0.655, on average, relative to those who do not. A possible reason for this could be correlated intermediary factors, such as that young people and those under 30 are more likely to be unemployed. This makes them eligible for legal assistance and services in many provinces that could positively impact their experience.

Self-represented litigants who reported having no education had a lower sentiment score, on average, by -0.297 relative to those with some education. This indicates that education plays a vital role in self-represented litigants' experiences of interaction with courts, which undoubtedly confirms that those with higher education find the experience to be better. The same is true for those with disabilities, who tended to have a lower sentiment score relative to those without disabilities, on average. This is intuitive as there has been a reported under-provision and inaccess to disability services that mitigate burdens particularly experienced by this group. Finally, those who took advantage of online services had a higher sentiment score, on average, of 0.140 relative to those who did not. This is interesting as it validates the move towards digital services, which could mitigate geographical barriers that currently disadvantage rural self-represented litigants. Further, access to online services could benefit single mothers, older persons, and those in precarious work situations who cannot commute or attend appointments at legal clinics and services.

However, the regression analysis does raise a few surprising findings. Notably, Indigenous and Asian self-represented litigants reported having a more positive experience self-representing than those from other ethnic groups. This contradicts existing literature and research, which widely suggests that Indigenous persons, in particular, are disproportionately disadvantaged by the justice system. However, upon further analysis, the model could suffer from selection bias as only 3 Indigenous persons were reported in the sample data. Therefore,

this finding may not be representative of the entire Indigenous self-represented population.

Further, on average, those who used a legal aid lawyer had a lower sentiment score of -0.191 relative to those who did not. This is surprising as it contradicts the finding in section 4.2 that legal aid lawyers had a positive prediction effect on self-represented litigants' experiences – which is also what the literature suggests. It is possible that the regression model may suffer from omitted variable bias as some factors are not included in the analysis that could be correlated with the variables of interest and outcome. An example of this could be, for example, unemployment, as described above. While we cannot extrapolate outside findings on these variables to our sample dataset, a few corrective measures could be undertaken, like specifying control variables in the model. Table 4 on the subsequent page regresses a model that includes control variables to account for these biases.

Table 4 shows the regression results for including control variables in our model. These control variables are, for example, case type, case location, virtual hearing, brought a support friend, opposing party was represented, and offered mediation. By including these variables in the analysis, we attempt to correct model inaccuracies caused by omitted variable bias. Interestingly, it appears as though our results change from the previous model specification. In this case, income, age, disability, legal aid lawyer, and online services became statistically insignificant. This means that the controlling regression model does not detect a relationship between these variables and the sentiment score, which our previous model did. Possible reasons for this could be that controlling for case characteristics diminished the impact of income, age, and disability on self-represented litigants' experiences because of clustering. By clustering similar cases together, the model incorrectly classifies the impact of these variables as random noise. Therefore, they are included in the error term but not reflected in our variable model.

Further research should focus specifically on these factors, as they are hypothesized to

Table 4: Regression Results (With Controls)

	Effects on Sentiment Score						
Female	-0.162*** (0.0570)						
No Education		-0.311*** (0.0962)					
Indigenous			0.338*** (0.0851)				
Asian				0.194* (0.104)			
Female X Income under $30k					-0.255*** (0.0644)		
Female X Black Ethnicity						-0.179*** (0.0563)	
Male X Black Ethnicity							0.144** (0.0565)
Observations	448	448	448	448	448	448	448

Standard errors in parentheses
* $p < .10$, ** $p < .05$, *** $p < .01$

be linked to self-represented litigants' experiences. More data should also be aggregated on a broader scale to encapsulate society-wide effects, such as unemployment, digitalization, and increases in disability benefits rates, on these variables. Having a legal aid lawyer and online service was most likely found insignificant in our Table 4 model due to multicollinearity issues with the control variables that characterize case type and assistance.

However, the results the model does produce regarding our existing variables of interest are still significant. Notably, the effect size of females having worse experiences than males stayed approximately the same. This presumes that controlling for case characteristics did not significantly alter the relationship between gender and sentiment. Females still report worse experiences relative to males, which comments on ongoing prejudicial biases against women in the courtroom. Further, the interaction term between females and low-income groups (below $30k) shows that these self-represented litigants are particularly disadvantaged, on top of the abovementioned effect. On average, these persons had a worse sentiment score, by -0.255 relative to males and higher-earning females. The analysis also suggests that black women have, on average, worse experiences than non-black women by a -0.179 lower sentiment score. This is particularly concerning as it could highlight prejudicial biases against women and vulnerable female groups. It also suggests that more attention should be devoted to these persons as they face the worst reported experiences of self-representing than any other gender group.

No education remained significant, correlating with a lower sentiment score, on average, by -0.311. Notably, this is higher than in our first model specification. This suggests that controlling for case characteristics exacerbated the effect size of low educational attainment. Indigenous and Asian effect sizes diminished, however, emphasizing that our inclusion of control variables may have partially corrected the omitted variable and selection bias. Notably, the interaction term between black persons and males is statistically significant

at a 5% level. On average, black males reported higher sentiment scores, compared to non-black males, by 0.144. This is surprising, as it seems to contradict the trend experienced by females. While this highlights that experiences differ substantially between female and male groups, it is possible that our model still suffers from uncorrected biases. As mentioned above, these could be correlated with broader social and economic trends that require further analysis to mitigate omitted variables biases.

5. Conclusion

This report makes two significant contributions to existing literature documenting the self-representation problem. First, it summarizes current research to provide a holistic interpretation of the issues facing self-represented litigants and re-emphasizes who they are: there is not one type. It shows the contrast between the ideal of self-representation portrayed through government sources and Ontario's law society to that of reality. There is an increasing number of self-represented litigants who face a plethora of social, economic, and geographical barriers to accessing justice. Further, systemic biases and prejudicial treatment of certain social groups amongst legal professionals and systems marginalize vulnerable populations. There are persistent difficulties in navigating a cumbersome legal system that is primarily a cause of unreasonably high costs but also intersects with broader issues such as language barriers, inaccess to culturally relevant services, and discrimination. Finally, issues of vexatiousness lead to power imbalances in the courtroom that protracts inequalities stratified along gender and ethnic lines.

The second significant contribution of this research is that it quantifies the extent of socioeconomic and demographic considerations' effects on experiences of self-representing. Positive experiences appear to be correlated with higher income, age, and education levels, as well as access to relevant services and inexpensive lawyers. Negative experiences, however, disproportionately characterize ethnic and female persons and those in lower-income groups that do not meet the eligibility criteria for legal aid. In performing this analysis, this report adds empirical data to a previously empty dataspace which is of great importance to policymakers on better reforming the legal system. It also raises awareness of an issue that is publicly less known and, through a data-oriented approach, suggests areas for future focus.

Moving forward, this data is of great use to the NSRLP for multiple reasons. First, it aligns with much of the already outputted research and contributes to those sources by

quantifying the extent of the self-representation problem. Secondly, it informs the NSRLP of who their respondents are and where their community base is primarily located. This implies that the organization could strengthen their outreach in areas with fewer respondents, or instead focus resources in Ontario to accommodate the majority of community persons. Third, by introducing quantitative and statistically rigorous analyses into a previously qualitative field, the NSRLP is positioned to lead the movement for a data- and evidence-oriented approach to addressing injustices facing self-represented litigants. More data needs to be tracked on who Canada's self-represented litigants are, the problems they face, and how those problems manifest outside of the legal system.

The reality of self-representation strays far from the ideal. To create a better system for future generations, these issues must be recognized today to improve tomorrow. It is negligent to assume they will be rectified overnight, but it is equally negligent to assume they are beyond correcting. The federal and provincial governments need to recognize the injustices that flow from effectuating one's right to self-represent, especially when that right may be the last resort for many seeking legal redress. It directly intersects with the Department of Justice's mission to achieve a fair, efficient, and accessible legal system and the government's efforts to reconcile with and accommodate vulnerable populations. As such, self-representation is inherently connected to other social justice issues that impact far more than the individual person, and it will take far more than that to motivate effective change.

6. References List

Access to Justice Task Force. (2009, September 1). *A Strategic Framework for Access to Justice in the Federal Civil Justice System*. Attorney-General's Department; Government of Australia. https://www.ag.gov.au/legal-system/publications/strategic-framework-access-justice-federal-civil-justice-system

Action Committee on Access to Justice in Civil and Family Matters. (2013). Access to Civil & Family Justice: A Roadmap for Change. *Canadian Forum on Civil Justice*, *58*. https://digitalcommons.osgoode.yorku.ca/cfcj/58

Arshad, A. (2007). A Self-Represented Family Litigant. *Canadian Forum on Civil Justice*, *1*(10), 12–15. http://cfcj-fcjc.org/sites/default/files/docs/2007/newsviews10-en.pdf#page=12

Baral, S., Bond, A., Boozary, A., Bruketa, E., Elmi, N., Freiheit, D., Ghosh, S. M., Goyer, M. E., Orkin, A. M., Patel, J., Richter, T., Robertson, A., Sutherland, C., Svoboda, T., Turnbull, J., Wong, A., & Zhu, A. (2021). Seeking shelter: homelessness and COVID-19. *FACETS*, *6*, 925–958. https://doi.org/10.1139/facets-2021-0004

Birnbaum, R., & Bertrand, N. B. (2013). THE RISE OF SELF-REPRESENTATION IN CANADA'S FAMILY COURTS: THE COMPLEX PICTURE REVEALED IN SURVEYS OF JUDGES, LAWYERS AND LITIGANTS. *The Canadian Bar Review*, *91*(1). https://cbr.cba.org/index.php/cbr/article/view/4288

Boyd, J.-P. E., & Bertrand, L. D. (2014, June). *Self-Represented Litigants in Family Law Disputes: Contrasting the Views of Alberta Family Law Lawyers and Judges of the Alberta Court of Queen's Bench*. CanLii. https://canlii.ca/t/285t

Boyd, J.-P. E., Bertrand, L. D., & Paetsch, J. J. (2014, April). *Self-represented Litigants in Family Law Disputes: Views of Judges of the Alberta Court of Queen's Bench*. CanLii. https://canlii.ca/t/285v

Brant, J. (2020, May 1). *Racial Segregation of Indigenous Peoples in Canada*. The Canadian Encyclopedia. https://www.thecanadianencyclopedia.ca/en/article/racial-segregation-of-indigenous-peoples-in-canada

British Columbia (Attorney General) *v.* Christie, [2007] 1 S.C.R. 873, 2007 SCC 21

Campbell, M., & Macfarlane, J. (2019). *The Self-Represented Litigants Case Law Database Occasional Research Series Self-Represented Litigants & Legal Doctrines of "Vexatiousness" An Interim Report from The National Self-Represented Litigants Project*. https://representingyourselfcanada.com/wp-content/uploads/2019/12/Vexatious-Litigant-Report-Final.pdf

Canadian Charter of Rights and Freedoms, s 7, Part 1 of the *Constitution Act*, 1982, being Schedule B to the *Canada Act 1982* (UK), 1982, c 11.

Canadian Judicial Council. (2006). *Statement of Principles on Self-represented Litigants and Accused Persons Adopted by the Canadian Judicial Council*. https://cjc-ccm.ca/sites/default/files/documents/2020/Final-Statement-of-Principles-SRL.pdf

Canadian Judicial Council, & McLachlin, B. (2004). *Ethical Principles for Judges*. https://cjc-ccm.ca/cmslib/general/news_pub_judicialconduct_Principles_en.pdf (Original work published 1988)

Cerniglia, C. (2020). The Civil Self-Representation Crisis: The Need for More Data and Less Complacency. *Georgetown Journal on Poverty Law and Policy*, *XXVII*(3). https://www.law.georgetown.edu/poverty-journal/wp-content/uploads/sites/25/2020/06/01-Cerniglia_Final_Proof.pdf

C.L.M. v M.J.S., 2017 BCSC 799 (CanLII), <https://canlii.ca/t/h3s5q>

Cooter, R., & Ulen, T. (2016). *Berkeley Law Berkeley Law Scholarship Repository Law and Economics, 6th edition*. http://www.econ.jku.at/t3/staff/winterebmer/teaching/law_economics/ss19/6th_edition.pdf

Currie, A. (2007). Civil Justice Problems and the Disability and Health Status of Canadians. *Journal of Law and Social Policy*, *21*(1), 31–48. https://digitalcommons.osgoode.yorku.ca/jlsp/vol21/iss1/2

Currie, A. (2009, May 12). *The Legal Problems of Everyday Life--The Nature, Extent and Consequences of Justiciable Problems Experienced by Canadians*. Www.justice.gc.ca; Department of Justice, Government of Canada. https://www.justice.gc.ca/eng/rp-pr/csj-sjc/jsp-sjp/rr07_la1-rr07_aj1/index.html

Day, A., Jones, R., Nakata, M., & McDermott, D. (2012). Indigenous Family Violence: An Attempt to Understand the Problems and Inform Appropriate and Effective Responses to Criminal Justice System Intervention. *Psychiatry, Psychology and Law*, *19*(1), 104–117. https://doi.org/10.1080/13218719.2010.543754

Department of Justice. (2018). *Canada's Department of Justice*. Justice.gc.ca; Government of Canada. https://www.justice.gc.ca/eng/rp-pr/cp-pm/about-aprop/index.html

Department of Justice. (2021). *Serious Legal Problems faced by Lesbian, Gay, Bisexual, and Other Sexual-Minority People in Western Canada: A Qualitative Study*. Government of Canada. https://www.justice.gc.ca/eng/rp-pr/jr/lgbtq/docs/rsd_rr2021_lgb-people-in-western-canada-eng.pdf

De Silva v. Fraser Health Authority, 2013 BCSC 659 (CanLII), <https://canlii.ca/t/fx2xz>

Devry Smith Frank LLP. (2017, November 2). *Increase In Self-Represented Litigants Increases Issues - Crime - Canada*. Www.mondaq.com. https://www.mondaq.com/canada/crime/642364/increase-in-self-represented-litigants-increases-issues

Dujardin v. Dujardin, 2018 ONCA 597 (CanLII), <https://canlii.ca/t/hv65x>

Eldridge *v*. British Columbia (Attorney General), [1997] 3 S.C.R. 624

Farrow, T. C. W., Lowe, D., Albrecht, B., Manweiller, H., & Simmons, M. E. (2012, March 27). *Addressing the Needs of Self-Represented Litigants in the Canadian Justice System: A White Papers Prepared for the Association of Canadian Court Administrators*. Association of Canadian Court Administrators. http://www.cfcj-fcjc.org/sites/default/files/docs/2013/Addressing%20the%20Needs%20of%20SRLs%20ACCA%20White%20Paper%20March%202012%20Final%20Revised%20Version.pdf

Feldstein, A. (2016). *The good, the bad and the ugly of self-representation* (Vol. 2). Feldstein Family Law Group. https://stream1.newswire.ca/media/2016/07/05/20160705_C1063_PDF_EN_726919.pdf

Gander, L., Lowe, D., & Stratton, M. (2005). The Civil Justice System and the Public:

    Highlights of the Alberta Pilot. *Alberta Law Review*, *42*(3), 803.

    https://doi.org/10.29173/alr1285

Girao v. Cunningham, 2020 ONCA 260 (CanLII), <https://canlii.ca/t/j6l6p>

Greacen, J. M. (2001). *Legal Information vs. Legal Advice: Developments during the last five*

    years.

    https://www.srln.org/system/files/attachments/Greacen%20ARTICLE%20Legal%20I

    nformation%20v%20Legal%20Advice%202001.pdf

Greacen, J. M. (2014). Self-Represented Litigants, The Courts, and the Legal Profession:

    Myths and Realities. *Family Court Review*, *52*(4), 662–669.

    https://doi.org/10.1111/fcre.12118

Gunn, B. L. (2016). *"Ignored to Death: Systemic Racism in the Canadian Healthcare*

    System. Office of the High Commissioner of Human Rights; United Nations.

    https://www.ohchr.org/sites/default/files/Documents/Issues/IPeoples/EMRIP/Health/

    UniversityManitoba.pdf

Gutierrez, L., Chadwick, N., & Wanamaker, K. A. (2018). Culturally Relevant Programming

    versus the Status Quo: A Meta-analytic Review of the Effectiveness of Treatment of

    Indigenous Offenders. *Canadian Journal of Criminology and Criminal Justice*, *60*(3),

    321–353. https://doi.org/10.3138/cjccj.2017-0020.r2

Hannaford-Agor, P., & Mott, N. (2003). Research on Self-Represented Litigation:

    Preliminary Results and Methodological Considerations. *The Justice System*

    Journal, *24*(2), 163–181. https://www.jstor.org/stable/27977139

Hann, R. G., Nuffield, J., Meredith, C., & Svoboda, M. (2002, September). *Court Site Study*

    of Adult Unrepresented Accused in the Provincial Criminal Courts. Department of

    Justice; Government of Canada. https://www.justice.gc.ca/eng/rp-pr/csj-sjc/ccs-

    ajc/rr03_la3-rr03_aj3/rr03_la3.pdf

Hok v Alberta, 2016 ABQB 651 (CanLII), <https://canlii.ca/t/gvqsq>

IBM. (2022). *What is a Decision Tree*. Www.ibm.com.

   https://www.ibm.com/topics/decision-trees

Jonsson v Lymer, 2020 ABCA 167 (CanLII), <https://canlii.ca/t/j6wg0>

Jukier, R. (2015). The Impact of Legal Traditions on Quebec Procedural Law: Lessons From
Quebec's New Code of Civil Procedure. *Canadian Bar Review*, *93*(1), 211–250.
https://canlii.ca/t/76v

Lieb, G. (2010, July 21). *Alberta Legal Services Mapping Project: Report for the Edmonton
Judicial District*. Canadian Forum on Civil Justice. http://cfcj-
fcjc.org/sites/default/files/docs/2010/mapping-edmonton-en.pdf

Loppie, S., Reading, C., & de Leeuw, S. (1986). INDIGENOUS EXPERIENCES WITH
RACISM AND ITS IMPACTS. *Clinical & Experimental Allergy*, *16*(S1), 9–9.
https://doi.org/10.1111/j.1365-2222.1986.tb03302.x

Lynch, J., & Davis, M. M. (2021, May 18). *What can be done about upward trend of self-
represented litigants - Law360 Canada*. Www.thelawyersdaily.ca.
https://www.law360.ca/articles/26793/what-can-be-done-about-upward-trend-of-self-
represented-litigants

Macfarlane, J. (2013, May). *The National Self-Represented Litigants Project: Identifying and
Meeting the Needs of Self-Represented Litigants*.
Www.representingyourselfcanada.com; National Self-Represented Litigants Project.
https://representingyourselfcanada.com/wp-content/uploads/2016/09/srlreportfinal.pdf

Macfarlane, J. (2021). *Critical Judicial Decisions for Self-Represented Litigants - Using
important case law that establishes rights for self-represented litigants*. CanLii;
National Self-Represented Litigants Project. https://canlii.ca/t/t3r7

McLachlin, B. (2008, September 4). *The Challenges We Face: Remarks of the Right
Honourable Beverley McLachlin, P.C. Chief Justice of Canada*. Www.scc-Csc.ca;
Supreme Court of Canada. https://www.scc-csc.ca/judges-juges/spe-dis/bm-2007-03-
08-eng.aspx

Naidu, A., & Macfarlane, J. (2021, March). *Pintea v Johns: An Updated Commentary*.
Representingyourselfcanada.com; National Self-Represented Litigants Project.
https://representingyourselfcanada.com/wp-content/uploads/2021/03/Pintea-Report-
2021-Final.pdf

Noone, M. A. (2009). Towards Integrated Legal Service Delivery. *SSRN Electronic Journal*.
https://doi.org/10.2139/ssrn.1425099

Ottewell v. Ottewell, 2013 ONSC 721 (CanLII), <https://canlii.ca/t/fw1gk>

Patrick, C. (2014). *Aboriginal homelessness in Canada : a literature review*. Canadian
Homelessness Research Network Press.

Pintea *v.* Johns, 2017 SCC 23, [2017] 1 S.C.R. 470

Reading, C., & Wien, F. (2009). *Health Inequalities and Social Determinants of Aboriginal
Peoples' Health*. https://www.ccnsa-nccah.ca/docs/determinants/RPT-
HealthInequalities-Reading-Wien-EN.pdf

Research and Statistics Division, D. of J. (2013, December 5). *JustFacts - Self-Represented
Litigants in Family Law*. Www.justice.gc.ca; Government of Canada.
https://www.justice.gc.ca/eng/rp-pr/fl-lf/divorce/jf-pf/srl-pnr.html

Royal Commission on Aboriginal Peoples. (1996). *Report of the Royal Commission on
Aboriginal Peoples*. Library and Archives Canada; Government of Canada.
https://www.bac-lac.gc.ca/eng/discover/aboriginal-heritage/royal-commission-
aboriginal-peoples/Pages/final-report.aspx

Select Standing Committee on Children and Youth. (2009). *Health and Well-Being of
Children in Care: Youth Justice Experiences and Outcomes Kids, Crime and Care
Joint Special Report*. https://cwrp.ca/sites/default/files/publications/en/BC-
YouthJusticeReport.pdf

Semple, N. (2010). *Cost-Benefit Analysis of Family Service Delivery: Disease, Prevention,
and Treatment*. Papers.ssrn.com. https://ssrn.com/abstract=3006813

Shushani, S., & Macfarlane, J. (2018, May). *When Judges See SRLs, Do They See Gender?
Observations on Gendered Characterizations in Judgments*.
Www.representingyourselfcanada.com; National Self-Represented Litigants Project.
https://representingyourselfcanada.com/wp-content/uploads/2018/05/Gendered-
Language-Paper-Final.pdf

Sinha, M. (2014). *Parenting and Child Support After Separation or Divorce*. Social and

> Aboriginal Statistics Division, Statistics Canada.

> https://www150.statcan.gc.ca/n1/pub/89-652-x/89-652-x2014001-eng.pdf

S.L.M.D. v. A.V.D., 2017 BCSC 394 (CanLII), <https://canlii.ca/t/h0rdp>

Stratton, M. (2011). *Alberta Legal Services Mapping Project: An Overview of Findings from

> the Eleven Judicial Districts*. Canadian Forum on Civil Justice. http://www.cfcj-

> fcjc.org/sites/default/files/docs/2011/mapping-final-en.pdf

Stratton, M., & Anderson, T. (2006). *Social, Economic and Health Problems Associated with

> a Lack of Access to the Courts A special report from the Civil Justice System and the

> Public Project*. http://cfcj-fcjc.org/sites/default/files/docs/2008/cjsp-socialproblems-

> en.pdf

Sullivan, C., & Macfarlane, J. (2021). *Tracking the Trends of the Self-Represented Litigant

> Phenomenon: Data from the National Self- Represented Litigants Project, 2019-2021*.

> www.representingyourselfcanada.com. https://representingyourselfcanada.com/wp-

> content/uploads/2021/10/Intake-Report-2021-FINAL.pdf

Sutherland, S., & Richards, C. (2017, November 1). *Supreme Court of Canada Endorses A

> New Approach to Self-Represented Litigants*. Www.lawnow.org.

> https://www.lawnow.org/supreme-court-of-canada-endorses-a-new-approach-to-self-

> represented-litigants/

Ward, R. (2019, February 22). *Building Trust Before Truth: How Non-Indigenous Canadians

> Become Allies*. Www.animikii.com. https://animikii.com/news/building-trust-before-

> truth-how-non-indigenous-canadians-become-allies

Zeidler, D. (2011). Building A Relationship: Perspectives From One First Nations

> Community Nouer une relation : le point de vue d'une communauté des premières

> nations. *Canadian Journal of Speech-Language Pathology and Audiology*, *35*(2).

> https://med-fom-osot-inclusive-

> campus.sites.olt.ubc.ca/files/2021/11/Zeidler_CJSLPA_2011.pdf

Zhu, Y., Yuan, Y., Gu, J., & Fu, Q. (2021). Neoliberalization and inequality: disparities in access to affordable housing in urban Canada 1981–2016. *Housing Studies*, 1–28. https://doi.org/10.1080/02673037.2021.2004093

7. Appendix A
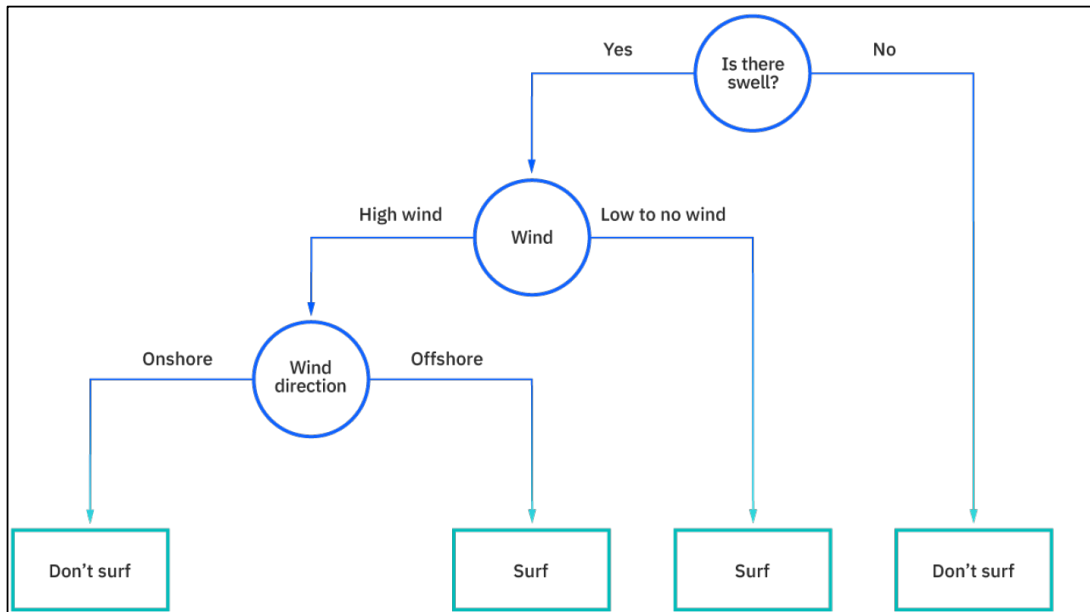
Above is an example of a decision tree for whether or not you should surf. As is evident, the branches each represent a decision that lead to the 'leaf' of the tree, i.e., whether the conditions are good for surfing on a particular day. The picture is taken from IBM (2022) available in the references list.

8. Appendix B

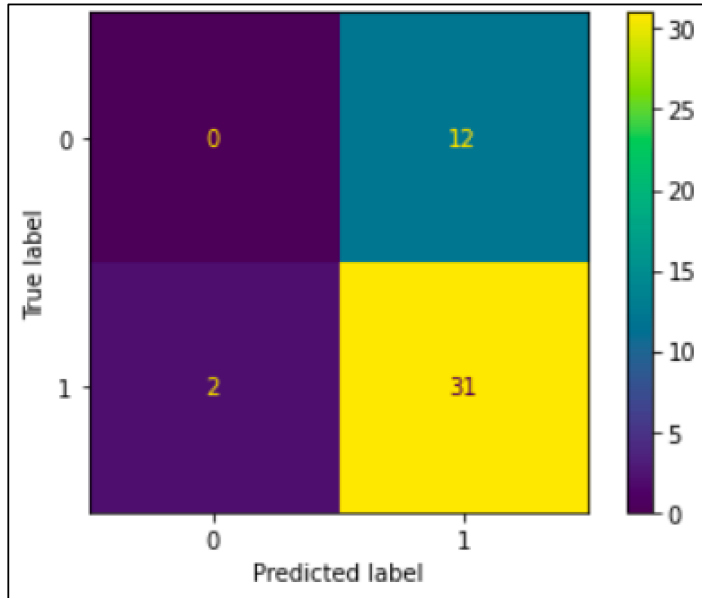*Logistic Regression Classifier Decision Boundary*
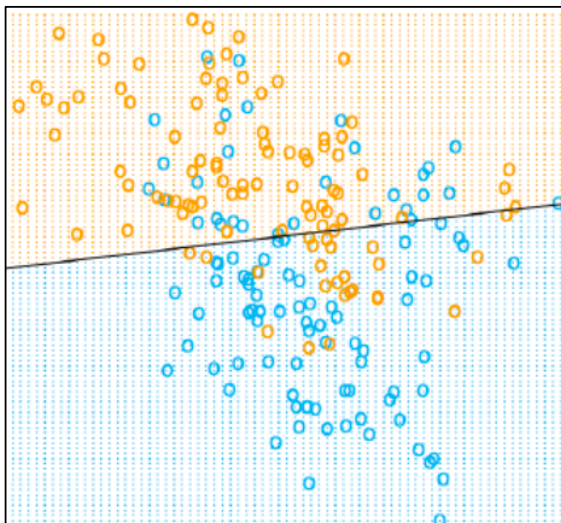
KNN Classifier Decision Boundary
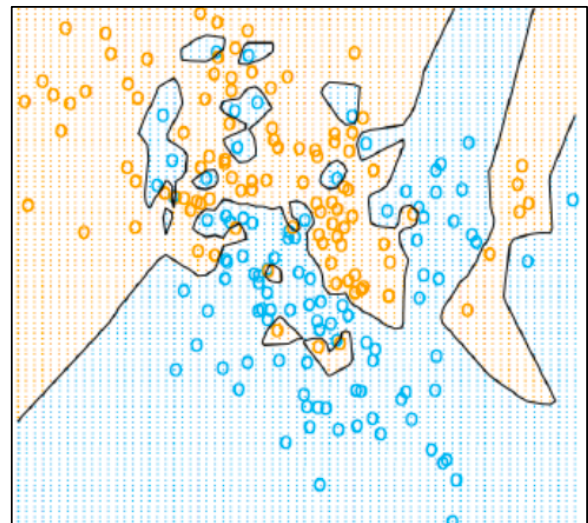
As shown above, the logistic regression classifier has a straight-line decision boundary, whereas the KNN classifier's is discontinuous. Notably, the 'clustering effect' of the KNN classifier is evident as it splits the data up into neighbourhoods as opposed to two regions. This derives from the logistic regression classifier's parametric attributes, which assume that the underlying relationship is linear. However, the KNN classifier is non-parametric meaning it doesn't make any assumptions about the underlying distributions of our predictor and outcome variables. This usually means that the KNN classifier is more robust. However, in linear cases, the logistic regression outperforms the KNN classifier.

9. Appendix C

KNN Classifier Confusion Matrix

Logistic Regression Classifier Confusion Matrix