

## **Public Data Set Variable Definitions & Data Dictionary**

### **Variable Definitions**

The variables associated with the Tracking (In)Justice data set provide specific information about every recorded case. A description of each of the variables is provided below. For more complex categories such as gender and race, the Tracking (In)Justice project looks toward definitions and categories adopted by Statistics Canada as reflected in evolving classifications of information collected about population demographics in this country.

#### **Date**

The date reported on which the death occurred.

#### **Name**

The name as presented in the media or official report emanating from the case. In some instances, the name is changed if the preferred name of the victim is suggested by family members or loved ones.

#### **Age**

The reported age of the victim.

#### **Gender**

Sex, gender expression, and gender identity are distinct concepts. Statistics Canada, for example, distinguishes between and provides definitions of “sex of person” and “gender of person”. Sex of person refers to sex assigned at birth based on the reproductive system and other physical characteristics.<sup>1</sup> Gender of person refers to “the gender that a person internally feels ('gender identity' along the gender spectrum) and/or the gender a person publicly expresses ('gender expression') in their daily life.”<sup>2</sup> A person’s gender may differ from sex assigned at birth (male or female), can change over time, and may be different than what appears on legal documents.

Within the Tracking (In)Justice data set we attempt to include information about a person’s gender, as opposed to their sex. Because the data set relies primarily on media reports and official government sources, in many cases a person’s gender will have been determined by the implicated police service, watchdog, or legal body based on physical characteristics, and reported in the media as such. We recognize the complexity and problems associated with relying so heavily on police-reported information, and as such the potential for misgendering. At times the broader community will be a more accurate source of information for identifying an individual’s gender.

We have a form on the website for family members, community members, and loved ones to correct information if incorrect information associated with gender appears in this data set.

#### **Race**

Determining racial categories is a complex, subjective, and potentially problematic process. Race is a complex social category that has evolved over time in tandem with data collection

---

<sup>1</sup> <https://www23.statcan.gc.ca/imdb/p3Var.pl?Function=DEC&Id=410445>

<sup>2</sup> <https://www23.statcan.gc.ca/imdb/p3Var.pl?Function=DEC&Id=410445>

## **Public Data Set Variable Definitions & Data Dictionary**

practices and processes. Race is different from ethnic origin and religion, and race categories are distinct from geographical region. One of the primary objectives of this project is to investigate if and how police actions disproportionately impact racialized peoples. Racial categories reflect a general understanding of race as a social descriptor in Canada, and the categories used in Tracking (In)Justice are drawn from existing government categorizations (Statistics Canada; Ontario Human Rights Commission).

Because race is a social construct, any description of race must be subject to ongoing critique, reflection, and reinterpretation. We are committed to re-examining and modifying the database's use of racial categories over time in consultation with affected communities.

The Tracking (In)Justice database relies largely on media and official government reports to help determine an individual's race. Information about a person's race will therefore often be drawn from police reports, reflecting police perceptions of an individual's race. Media reports may also draw from interviews with community members or reports from independent oversight bodies. There may be discrepancies, therefore, between how an individual would have self-identified, and the race as assigned by others whether it be the police or the broader community. We have created a form for victims' family members and loved ones to submit corrections to this data set.

Despite the difficulties accessing self-reported racial data for police-involved deaths, the perceptions of police and broader societal perceptions of an individual's race are directly relevant to understanding the links between race and police-involved deaths. We therefore believe that the race information in this database is sufficiently reliable to sustain analysis.

Much of the racial data from 2000 to 2020 originates from the CBC's Deadly Force data set. The CBC data set used a condensed version of Statistics Canada visible minority groups.<sup>3</sup> CBC notes that the primary sources of information to fill this category were official police reports or reports from independent police watchdogs. CBC researchers also made best efforts to assign a race or ethnicity to each victim using a collection of family statements in media reports, interviews with friends or neighbours, pictures from social media or obituaries, and through logical inference (e.g., if a victim was a band member of a First Nation community, it was assumed they were Indigenous).

Tracking (In)Justice has decided not to independently assign race to an individual by looking at photographs or examining their name. However, if victims are reported as being a specific race, they are identified as such in our data set. These methods are imperfect and may result in a number of "unknown" instances.

### **Province**

The province or territory where the death occurred.

### **Police Service**

The police service that interacted with the individual who died. In some cases, more than one police service was present at the time of death. We list both police services if this was the case.

---

<sup>3</sup> <https://www23.statcan.gc.ca/imdb/p3Var.pl?Function=DEC&Id=45152>

## Public Data Set Variable Definitions & Data Dictionary

### Highest Level of Force

The highest level of force used by police during an incident where a fatality occurs. Police force was intentionally used in all the cases included in the data set.

In the majority of cases, the highest level of force used is gunshot(s). There are a few cases where police claim that they did not fire their weapons, or it is unclear what happened; these details are indicated for our team internally.

Although many individuals in the database did die as a direct result of the police use of force, confirming an official cause of death is beyond the scope of this project. As a result, inclusion in the database does not necessarily mean there is a causal link between the type of force used and the individual's death.

### Data Dictionary

Column Label	Description	Variable Type	Values
Name_First	First name of victim	Text	
Name_Middle	Middle name of victim (if any or available)	Text	
Name_Last	Last name of victim	Text	
Age	Age of the victim at death	Numerical	
Date	Date of death	Date	
Gender	Gender of victim	Categorical	Man, Woman, Non-binary, Two Spirit, Unknown
Race	Race of victim	Categorical	White, Black, Latin American, South Asian, Asian, Indigenous, Arab, Unknown
Province/Territory	Province or Territory that incident occurred in	Categorical	ON, QC, AB, etc.
Police_Service	Police Service involved in incident	Text	e.g.: RCMP, Halton Region Police Service
Highest_Level_Force	The highest level of police force used in the incident	Categorical	Gunshot, Restraint, Intermediate Weapon, Taser/CEW, Physical Force, Unknown, Other