CIMI

CODEBOOK

MARCH 2020

Institut canadien pour les identités et les migrations | The Canadian Institute for Identities and Migration

ACS • AEC
Association for Canadian Studies • Association d'études canadiennes

TABLE OF CONTENTS

STEP 1: RECODE VARIABLES

(Dataset: Census 1991, 1996, 2001, 2006, 2016 and NHS 2011.)

1. Filter

ORIGINAL VARIABLE	RECODED VARIABLE
AGE 2016: "AGECONT" (Continuous variable) 2011, 2006, 2001, 1996, 1991: "AGE" (Continuous variable)	"AGEfilter" (Age group from 18-64.) 18 thru 64 → 1 = "18 – 64 years old" ELSE → 0
LABOUR FORCE STATUS 2016, 2011, 2006, 2001, 1996, 1991: "LFTag" (Labour: Labour force status) 22/0/-3 = Not applicable, < 15 years 1 = Employed - Worked in reference week - Armed Forces 2 = Employed - Worked in reference week - Civilian 3 = Employed - Absent in reference week - Armed Forces 4 = Employed - Absent in reference week - Civilian 5 = Unemployed - Temporary layoff - Experienced - Did not look for work 6 = Unemployed - Temporary layoff - Experienced - Looked for full-time work 7 = Unemployed - Temporary layoff - Experienced - Looked for part-time work 8 = Unemployed - New job - Experienced - Did not look for work 9 = Unemployed - New job - Experienced - Looked for full-time work 10 = Unemployed - New job - Experienced - Looked for part-time work 11 = Unemployed - New job - Inexperienced - Did not look for work 12 = Unemployed - New job - Inexperienced -Looked for full-time work 13 = Unemployed - New job - Inexperienced - Looked for part-time work 14 = Unemployed - Looked for full-time work - Experienced 15 = Unemployed - Looked for part-time work - Experienced 16 = Unemployed - Looked for full-time work - Inexperienced 17 = Unemployed - Looked for part-time work - Inexperienced 18 = Not in the labour force - Last worked in 2011 19 = Not in the labour force - Last worked in 2010 20 = Not in the labour force - Last worked before 2010 21 = Not in the labour force - Never worked	"EMPL" (People who are currently employed - worked or absent in reference week.) 1 thru 4 → 1 = "Currently employed" 5 thru 21 → 0 22/0/-3 → Missing value

PAID WORKERS	"COWfilter" (Paid workers working for wages.)
2016, 2011: "COWD" (Labour: Class of worker (derived)	1 → 1 = "Paid workers"
-5 = Did not work in 2010	2,3,4,5,6 → 0
-3 = Not applicable, < 15 years	-5, -3 → Missing value
1 = Employee	
2 = Unpaid family workers	
3 = Self-employed, without paid help, incorporated	
4 = Self-employed, with paid help, incorporated	
5 = Self-employed, without paid help, unincorporated	
6 = Self-employed, with paid help, unincorporated	
2006, 2001, 1996: "COWD" (Labour: class of worker (derived)	"COWfilter" (Paid workers working for wages.)
1 = Unpaid family workers - Worked without pay for a relative in a family business or farm	5 → 1 = "Paid workers"
2 = Not applicable	1,2,3,4,6,7, → 0
3 = Paid worker - Originally self-employed without paid help, incorporated	
4 = Paid worker - Originally self-employed with paid help, incorporated	
5 = Paid Worker - Working for wages, salary, tips or commission	
6 = Self-employed without paid help, not incorporated	
7 = Self-employed with paid help, not incorporated	
1991: "COWD" (Labour: Class of worker (derived)	"COWfilter" (Paid workers working for wages.)
0 = Not applicable	2 → 1
1 = Unpaid family workers - Worked without pay for a relative in a family business or farm	0,1,3,4,5,6 → 0
2 = Paid Worker - Working for wages salary tips or commission	
3 = Paid worker - Originally self-employed without paid help incorporated	
4 = Paid worker - Originally self-employed with paid help incorporated	
5 = Self-employed without paid help not incorporated	
6 = Self-employed with paid help not incorporated	

2. Dependent Variables

ORIGINAL VARIABLE	RECODED VARIABLE
WAGES 2016, 2011, 2006, 2001, 1996, 1991: "WAGES"/ "WAGESP" (Income: wages and salaries): Continuous variable	"Wages" refers to gross wages and salaries before deductions for such items as income taxes, pension plan contributions and employment insurance premiums during the reference period. **(1) For descriptive table:** + In order to calculate the WAGE CPI adjusted, choose the base year (in this project, the base year is 2002, CPI = 100). Real Wage = Wage in year X* 100/ CPI in year X For example: CPI in 2011 in Toronto is 120 (https://www150.statcan.gc.ca/t1/tbl1/en/cv.action?pid=1810000501#timeframe). Average wage of non-immigrants in Toronto is $65,795. Then the real wage = $65,795*100/120= $54,829. **(2) For regression table:** + Create new variable named sqrtWAGES by going to Transform → Compute → sqrtWAGES → Oke Define the missing value for the new variable.
LICOs 1996, 1991: "**LoInc**"/ 2001: **LOINC**/ 2006: **LOINCB**/ 2016, 2011: **LoLICOB** (Low income status (person) -3 = Concept not applicable (Person not in an economic family < 15 years old) 1 = Non-low-income person (At or above LICO-BT) 2 = Low-income person (Below LICO-BT) 3 = Concept not applicable (Territories or reserves / at or above LICO-BT) 4 = Concept not applicable (Territories or reserves / below LICO-BT) → The categories vary slightly between years, but "2" always means "member of a low income economic family or low income person", so we recode in the same way for all data years. Note: LICO after tax is only available since 2006 (https://www12.statcan.gc.ca/census-recensement/2006/ref/dict/fam019a-eng.cfm). Therefore, LICO before tax is chosen to use in this project to ensure the consistency over years.	LICOre (Refers to the percentage of individuals who have lived below Statistics Canada's low income cut-offs, before-tax). 2 → 1 = Low-income person (below LICO-BT) 1 → 0 = Non-low-income person (at or above LICO-BT) -3, 3,4 → Missing value

LIM-BT 2016, 2011: "LoLIMB"(Income: Low income status based on LIM-BT) 1 = Non-low-income person (at or above LIM-BT) 2 = Low-income person (below LIM-BT) 3 = Concept not applicable (Territories or reserves / at or above LIM-BT) 4 = Concept not applicable (Territories or reserves / below LIM-BT) 1991, 1996, 2001, 2006: not available	LIMMIre (Refers to the percentage of individuals who have lived below Statistics Canada's low Income Measure-Market Income.) 2 → 1 = Low-income person (below LIM-MI) 1 → 0 = Non-low-income person (at or above LIM-MI) 3,4 → Missing value
LIM-MI 2016, 2011: "LOLIMMI" (Income: Low income status based on LIM-MI) 1 = Non-low-income person (at or above LIM-MI) 2 = Low-income person (below LIM-MI) 3 = Concept not applicable (territories or reserves / at or above LIM-MI) 4 = Concept not applicable (territories or reserves / below LIM-MI) 1991,1996, 2001, 2006: not available	LIMMIre (Refers to the percentage of individuals who have lived below Statistics Canada's low Income Measure-Market Income.) 2 → 1 = Low-income person (below LIM-MI) 1 → 0 = Non-low-income person (at or above LIM-MI) 3,4 → Missing value
MBM 2016, 2011: "LoMBM" (Income: Low income status based on MBM) -3 = Concept not applicable (person not in an economic family < 15 years old) 1 = Non-low-income person (at or above MBM) 2 = Low-income person (below MBM) 3 = Concept not applicable (reserves / at or above MBM) 4 = Concept not applicable (reserves / below MBM) 1991, 1996, 2001, 2006: not available	LOMBMre (Refers to the percentage of individuals who have lived below Human Resources and Skills Development Canada's Market Basket Measure.) 2 → 1 = Low-income person (below MBM) 1 → 0 = Non-low-income person (at or above MBM) -3,3,4 → Missing value
LABOUR FORCE PARTICIPATION 2016, 2011, 2006, 2001, 1996, 1991: "LFTag" (Labour: Labour force status) 22/0/-3 = Not applicable, < 15 years 1 = Employed - Worked in reference week - Armed Forces 2 = Employed - Worked in reference week - Civilian 3 = Employed - Absent in reference week - Armed Forces 4 = Employed - Absent in reference week - Civilian 5 = Unemployed - Temporary layoff - Experienced - Did not look for work 6 = Unemployed - Temporary layoff - Experienced - Looked for full-time work 7 = Unemployed - Temporary layoff - Experienced - Looked for part-time work 8 = Unemployed - New job - Experienced - Did not look for work 9 = Unemployed - New job - Experienced - Looked for full-time work 10 = Unemployed - New job - Experienced - Looked for part-time work	"LFACTre" (Refers to the percentage of individuals who are active in the labour force, either employed or unemployed - but looking for work.) 1 thru 17 → 1 = "In the labor force" 18 thru 21 → 0 = "Not in the labor force" 22/0/-3 → Missing value

11 = Unemployed - New job - Inexperienced - Did not look for work 12 = Unemployed - New job - Inexperienced - Looked for full-time work 13 = Unemployed - New job - Inexperienced - Looked for part-time work 14 = Unemployed - Looked for full-time work - Experienced 15 = Unemployed - Looked for part-time work - Experienced 16 = Unemployed - Looked for full-time work - Inexperienced 17 = Unemployed - Looked for part-time work – Inexperienced 18 = Not in the labour force - Last worked in 2011 19 = Not in the labour force - Last worked in 2010 20 = Not in the labour force - Last worked before 2010 21 = Not in the labour force - Never worked	
EMPLOYMENT RATE Employment rate (Derived from LFTag)	"EMPL" ((Employment/population ratio) Number of employed persons expressed as a percentage of the population 18-64 years.) 1 thru 4 → 1 = "Employed" 5 thru 21 → 0 = "Others" 22/0/-3 → Missing value
UNEMPLOYMENT RATE Unemployment rate (Derived from LFTag)	"UNEMPL" (Number of unemployed persons expressed as a percentage of the labour force.) 5 thru 17 → 1 = "Unemployed" 1 thru 4 → 0 = "Employed" 18 thru 21 → Missing value 22/0/-3 → Missing value
FULL TIME EMPLOYMENT RATE 2016, 2011, 2006, 2001, 1996, 1991: "FPTim" (Labour: Full-time or part-time weeks) -5 = Did not work in 2010 -3 = Not applicable, < 15 years 1 = Worked mainly full-time weeks in 2010 2 = Worked mainly part-time weeks in 2010	Already recoded for the filter. ("FULLTIME" (Refers to the percentage of individuals who are working full-time.) 1 → 1 = "Full time" 2 → 0 = "Part time" -5, -3 → Missing value
NON OFFICIAL LANGUAGE AT WORK 2016:"LNWADR"(Language: Language used at work (part A) – Derived -3 = Not applicable 1 = English 2 = French	NOLre (Refers to the percentage of individuals using a non-official language most often at work.) 3 thru 216 → 1 = "Non-official languages" ELSE → 0 = "Official language(s)" -3 → Missing value

3 to 216 = Other languages

217 = English and French

218 = English and non-official language

219 = French and non-official language

220 = English, French and non-official language

2011: "LnWADr" (Language of work (A) - Derived - Part A

1 = English

2 = French

3 to 191 = Other languages

192 = English and French

193 = English and non-official language

194 = French and non-official language

195 = English, French and non-official language

196 = Not applicable

NOLre (Non-official language at work.)

3 thru 191 → 1 = "Non-official languages"

ELSE → 0 = "Official language(s)"

196 → Missing value

2006: LNWADR: Language of work derived - Part A

1 = English

2 = French

3 to 128 = Other languages

129 = English and French

130 = English and Non-official language

131 = French and Non-official language

132 = English, French and Non-official language

133 to 151= Other languages

152 = Non Response

NOLre (Non-official language at work.)

3 thru 128 → 1 = "Non-official languages"

133 thru 151 → 1 = "Non-official languages"

ELSE → 0 = "Official language(s)"

152 → Missing value

2001: LNWADR: Language of work derived - Part A

1 = English

2 = French

3 to 78 = Other languages

79 = Not applicable (Population 15 years and over who have not worked since 2000)

80 to 126 = Other languages

127 = English and French

128 = English and Non-official language

129 = French and Non-official language

130 = English, French and Non-official language

NOLre (Non-official language at work.)

3 thru 78 → 1 = "Non-official languages"

80 thru 126 → 1 = "Non-official languages"

ELSE → 0 = "Official language(s)"

79 → Missing value

1991, 1996: Not available

SUBSIDY HOUSING 2016, 2011: "SUBSIDY" (Subsidized housing) 0 = "No, not a subsidized dwelling" 1 = "Yes, a subsidized dwelling" -3 = "Not applicable (i.e., farm dwelling, owner-occupied dwelling, band housing)" 1991, 1996, 2001, and 2006: Not available	SUBSIDYre ("Subsidy housing" refers to the percentage of renters who live in subsidized housing (i.e. rent applied to income, social housing, public housing, government-assisted housing, or non-profit housing.)) 0 → 0 = "No, not a subsidized dwelling" 1 → 1 = "Yes, a subsidized dwelling" -3 → Missing value

3. Key Independent Variables

ORIGINAL VARIABLE	RECODED VARIABLE
IMMIGRANT STATUS 2016, 2011: "ImmDer" (Immigrant status) 1 = "Non-immigrants" 2 = "Immigrants" 3 = "Non-permanent residents"	IMMre (Includes persons who are, or who have ever been, landed immigrants or permanent residents.) "IMMre" (Immigrants and non-immigrants) 1 → 0 = "Non-immigrants" 2 → 1 = "Immigrants" 3 → Missing value
2006, 2001, 1996: IMMDER: Immigrant status 1 = Not applicable (Institutional resident) 2 = Non-immigrants 3 = Immigrants 4 = Non-permanent residents	"IMMre" (Immigrants and non-immigrants.) 2 → 0 = "Non-immigrants" 3 → 1 = "Immigrants" 1,4 → Missing value
1991: ImmDer (Immigrant status) 1 = Non-immigrant population 2 = Immigrant population 3 = Non-permanent resident population 4 = Not applicable (Institutional resident)	"IMMre" (Immigrants and non-immigrants.) 1 → 0 = "Non-immigrants" 2 → 1 = "Immigrants" 3,4 → Missing value

YEAR OF IMMIGRATION 2016: YRIM (Immigration: Year of immigration)/ PERIMMA (by 5 year group) 1 =Non-immigrants 2 =Before 1981 3 =1981 to 1985 4 =1986 to 1990 5 =1991 to 1995 6 =1996 to 2000 7 =2001 to 2005 8 = 2006 to 2010 9 = 2011 to 2016 10 = Non-permanent residents 2011, 2006, 2001, 1996: YRIM/ 1991: ImmYear Immigration: Year of immigration (Continuous variable) -4 = Non-permanent residents -3 = Non-immigrants	YRIMre (Refers to the year in which the immigrant obtained landed immigrant status by immigration authorities.) For 2016: 9 → 1 = "Recent immigrants" 2 to 8 → 2 = "Established immigrants" 1,10 → Missing value YRIMre 2006 - 2011 → 1 = "Recent immigrants" …(Before) → 0 = "Established immigrants" -3, -4 → Missing value ….
ADMISSION CATEGORIES 2016: IMMCAT5 (Immigration: Admission category and applicant type) 1 = Non-immigrants 2 = Immigrants who landed before 1980 3 = Non-permanent residents 100010 = Economic immigrants - Principal applicants 100020 = Economic immigrants - Secondary applicants 200000 = Immigrants sponsored by family 300000 = Refugees 400000 = Other immigrants Other years: Not available	IMMCATre ('Admission category' refers to the name of the immigration program or group of programs under which an immigrant has been granted for the first time the right to live in Canada permanently by immigration authorities.) 100010 → 1= Economic immigrants - Principal applicants 100020 → 2 = Economic immigrants - Secondary applicants 200000 → 3 = Immigrants sponsored by family 300000 → 4 = Refugees 400000 → 5 = Other immigrants 1,2,3 → Missing value

PROVINCE

2016, 2011, 2006, 2001, 1996, 1991: "PR"
(Province or territory of current residence)

10 = Newfoundland and Labrador

11 = Prince Edward Island

12 = Nova Scotia

13 = New Brunswick

24 = Quebec

35 = Ontario

46 = Manitoba

47 = Saskatchewan

48 = Alberta

59 = British Columbia

60 = Yukon

CMA

2016, 2011, 2006, 2001, 1996, 1991, "CMA"
(Census metropolitan area of current residence)

PROVINCES	CITIES (CMAS)
Newfoundland	St. John's
Prince Edward Island	—
Nova Scotia	Halifax
New Brunswick	Moncton, Saint John
Québec	Montréal, Québec, Saguenay, Sherbrooke, Trois-Rivières
Ontario	Barrie, Brantford, Guelph, Hamilton, Kingston, Belleville, Kitchener-Cambridge-Waterloo, London, Oshawa, Ottawa-Gatineau, Peterborough, St. Catharines-Niagara, Greater Sudbury, Thunder Bay, Toronto, Windsor
Manitoba	Winnipeg
Saskatchewan	Regina, Saskatoon
Alberta	Edmonton, Calgary, Lethbridge
British Columbia	Abbotsford, Kelowna, Vancouver, Victoria

(1) For descriptive table: Do not need to recode.

(2) For regression:

Step 1: Create dummy variable for each province/CMA.

Ex: QCre (People from Quebec)

24 → 1 = "Quebec"

ELSE → 0 = "Others"

ONre (People from Ontario)

35 → 1 = "Ontario"

ElSE → 0 = "Others"

Do the same for all 10 provinces and all 35 CMAs

Step 2: Create new variables showing the interaction between immigrant status and geography (Immigrant Status X Geography).

Ex: "NFxIMM", "QCxIMM" …

4. Control Variables

ORIGINAL VARIABLE	RECODED VARIABLE
VISIBLE MINORITY 2016: DVISMIN (same as 2011, but include "14 = Aboriginal peoples") 2011, 2006, 2001, 1996: DVisMin (Population group: visible minority status) 1 = Chinese 2 = South Asian 3 = Black 4 = Arab/West Asian 5 = Filipino 6 = Southeast Asian 7 = Latin American 8 = Japanese 9 = Korean 10 = Pacific Islander 11 = Visible minority, n.i.e. 12 = Multiple visible minority 13 = Other 14/15 = Not Applicable	"VISMINre" (Refers to persons who are non-Caucasian in race or non-white in colour and who do not report being Aboriginal.) 2016: 1 thru 12 → 1 = "Visible minority" 13 thru 14 → 0 = "Not visible minority" 2011, 2006, 2001, 1996: "VISMINre" (Visible minority) 1 thru 12 → 1 = "Visible minority" 13 → 0 = "Not visible minority" 14/15 → Missing value
1991: VisMin (Visible Minority Status) 1 = Blacks 2 = South Asians 3 = Chinese 4 = Koreans 5 = Japanese 6 = Southeast Asians 7 = Filipinos 8 = Other Pacific Islanders 9 = West Asians and Arabs 10 = Latin Americans 11 = Multiple Visible Minorities 12 = Not a Member of a Visible Minority 13 = Not applicable	"VISMINre" (Visible minority) 1 thru 11 → 1 = "Visible minority" 12 → 0 = "Not visible minority" 13 → Missing value *A note from Statscan: To obtain an overall estimate for "Not a visible minority", sum the following categories: 'Other (not a visible minority)' and 'Aboriginal peoples'.

SEX All years: "SEX" (sex) 1 = "Female" 2 = "Male"	SEXre 1 → 1 = "Female" 2 → 0 = "Male"
AGE 2016: AGECONT (Continuous variable) 2011, 2006, 2001, 1996, 1991: "AGE" (Continuous variable)	**(1) For descriptive table:** "AGEre" (CIMI age groups) 18 thru 24 → 1 = "18-24 years old" 25 thru 44 → 2 = "25-44 years old" 45 thru 64 → 3 = "45-64 years old" MISSING → SYSMIS **(2) For regression:** Use the original var (AGECONT).
HIGHEST LEVEL OF EDUCATION 2016, 2011, 2006: "HCDD" -3/14 = Not applicable (< 15 years) 1 = No certificate, diploma or degree 2 = High school diploma or equivalency certificate 3 = Other trades certificate or diploma 4 = Registered apprenticeship certificate 5 = College, CEGEP or other non-university certificate or diploma from a program of 3 months to less than 1 year 6 = College, CEGEP or other non-university certificate or diploma from a program of 1 year to 2 years 7 = College, CEGEP or other non-university certificate or diploma from a program of more than 2 years 8 = University certificate or diploma below bachelor level 9 = Bachelor's degree 10 = University certificate or diploma above bachelor level 11 = Degree in medicine, dentistry, veterinary medicine or optometry 12 = Master's degree 13 = Earned doctorate degree	EDUre (Education: Highest certificate, diploma or degree.) **(1) For descriptive table:** "EDUre" 1 thru 8 → 0 = "< Bachelor's degree" 9 thru 13 → 1 = "Bachelors degree and above" -3/14 → Missing value
2001, 1996, 1991: DGREER (Highest degree, certificate or diploma) 1 = None 2 = Secondary (high) school graduation 3 = Trades certificate or diploma 4 = Non-university certificate or diploma	"EDURe" 1 thru 5 → 0 = "< Bachelor's degree" 6 thru 10 → 1 = "Bachelors degree and above" 0,11 → Missing value

5 = University certificate or diploma below bachelor

6 = Bachelor's degree(s)

7 = University certificate or diploma above bachelor

8 = Degree in medicine, dentistry, veterinary medicine, or optometry

9 = Master's degree(s)

10 = Earned doctorate degree

0/11 = Not applicable

(2) For regression: Use the original education variable (note that the variable should be in the order from low to high level of education).

KNOWLEDGE OF OFFICIAL LANGUAGES

2016, 2011, 2006, 2001, 1991: "OLN"

1 = "English only"

2 = "French only"

3 = "Both English and French"

4 = "Neither English nor French"

0/5 = "Not applicable"

1996: OLN

6 = Not applicable

1 = Non-response

2 = English only

3 = French only

4 = Both English and French

5 = Neither English nor French

KOLre (Knowledge of official languages: Refers to the ability to conduct a conversation in English only, in French only, in both English and French or in none of the official languages of Canada).

(1) For descriptive table: Do not need to recode.

"KOLre" (Knowledge of Official Languages)

(0/ 5 → Missing value)

KOLre (Knowledge of official language)

2 → 1 = English only

3 → 2 = French only

4 → 3 = Both English and French

5 → 4 = Neither English nor French

6, 1 → Missing value

(2) For regression: Englishre
(Conduct a conversation in English only)

1 → 1 = "English only"

2 thru 4 → 0 = "Others"

Frenchre (Conduct a conversation in French only)

2 → 1 = "French only"

1 → 0 = "Others"

3thru 4 → 0 = "Others"

BothEFre (conduct a conversation in both English and French)

3 → 1 = "Both English and French"

1thru 2 → 0 = "Others"

4 → 0 = "Others"

	NeitherEFre (conduct a conversation in neither English or French) 4 → 1 = "Neither English nor French" 1thru 3 → 0 = "Others"
OCCUPATION 2016: NOC16BRD -5 = Did not work in 2015 and 2016 -3 = Not applicable, < 15 years 1 = Management occupations 2 = Business, finance and administration occupations 3 = Natural and applied sciences and related occupations 4 = Health occupations 5 = Occupations in education, law and social, community and government services 6 = Occupations in art, culture, recreation and sport 7 =Sales and service occupations 8 = Trades, transport and equipment operators and related occupations 9 =Natural resources, agriculture and related production occupations 10 = Occupations in manufacturing and utilities 2011: NOC11BRD 2006: NOCSMAJ …	**(1) For descriptive table:** "MGMT" (Management occupations) 1 → 1 = Management occupations 2 thru 10 → 0 = Others 11/-3 → Missing value **(2) For regression:** - For all indicators except for "Subsidy" and "LICO": Create dummy variables for NOC16BRD (original variable). Use these dummy variables (from OCC_2 → OCC_10) and OCC_1 (managers) as ref occupation. - For "Subsidy" and "LICO" indicators: + NOC 1 → 1 = Management occupations 2 → 2 = Business, finance and administration occupations 3 → 3 = Natural and applied sciences and related occupations 4 → 4 = Health occupations 5 → 5 = Occupations in education, law and social, community and government services 6 → 6 = Occupations in art, culture, recreation and sport 7 → 7 = Sales and service occupations 8 → 8 = Trades, transport and equipment operators and related occupations 9 → 9 = Natural resources, agriculture and related production occupations 10 → 10 = Occupations in manufacturing and utilities -5 → 11 = "Did not work in 2015 and 2016" -3 → Missing value + Create dummy variable for NOC (recoded variable). Use these dummy variables (from NOC_2 → NOC_11) and NOC_1 (managers) as ref occupation.

FULL-TIME/PART-TIME WORKERS 2016, 2011, 2006, 2001, 1996, 1991: "FPTim" (Labour: Full-time or part-time weeks worked) -5 = Did not work -3 = Not applicable, < 15 years 1 = Worked mainly full-time weeks 2 = Worked mainly part-time weeks	**(1) For descriptive table:** Don't need to recode. **(2) For regression:** "FULLTIME" (People who work mainly full-time) 1 → 1 = "Full time" 2 → 0 = "Part-time" -5, -3 → Missing value Note: This is a control for only "Wages" indicator.
MOBILITY STATUS MOB5PR" (Mobility 5: Mobility status - Place of residence 5 years ago) -3 = Not applicable - age exclusion 1 = Non-mover 2 = Non-migrant 3 = Intraprovincial migrant 4 = Interprovincial migrant 5 = External migrant	**(1) For descriptive table:** Don't need to recode. **(2) For regression:** MOB5re (Mobility 5: Mover and Non-mover) 1 thru 2 → 0 = "Non-movers" 3 thru 5 → 1 = "Movers" -3 → Missing value

STEP 2: FILL THE DESCRIPTIVE TABLES & T-TEST

A. FOR CONTROL VARIABLES

a. Turn on the filter → AGEPfilter = 1 (Age 18 – 64);

b. Analyze → Descriptive Statistics → Crosstab → Rows= PR/CMA (enter one at a time), Column= all control vars, Layer 1= IMMre.

B. FOR INDICATORS

1. Wages (full-time)

a. Turn on the filter → (AGEPfilter = 1) & (EMPL = 1) & (FULLTIME = 1) & (COWfilter = 1);

b. Analyze → Compare mean → Dependent var: wages; Layer 1: IMMre, Layer 2: PR/CMA (enter one at a time);

c. For t-test (don't change the filter) —2 steps:

* Split file – Group based on PR/CMA – Oke

* Compare means → independent sample t-test: test var – WAGES; Group variable: IMMre (define g1: 0 (Non-Imm); g2: 1 (Imm) → take p value from independent sample test → t-test for equality of means (Sign(2-tailed) → Note * (p<.05)/ ** (p<.01/ *** (p<.001) on the absolute difference between Imm and Non-imm.

d. Turn off "split file" function.

2. Wages (part-time)

a. Turn on the filter → (AGEPfilter = 1) & (EMPL = 1) & (PARTTIME = 1) & (COWfilter = 1));

b. Analyze → Compare mean → Dependent var: wage; Layer 1: IMMre, Layer 2: PR/CMA (enter one at a time);

c. For t-test (don't change the filter) —2 steps:

* Split file – Group based on PR/CMA – Oke

* Compare means → independent sample t-test: test var – WAGES; Group variable: IMMre (define g1: 0 (Non-Imm); g2: 1 (imm) → take p value from independent sample test → t-test for equality of means (Sign(2-tailed) → Note * (p<.05)/ ** (p<.01/ *** (p<.001) on the absolute difference btw imm and non-imm.

d. Turn off "split file" function.

3. LICOs, Labour force participation, Employment rate, Unemployment rate, Full-time employment rate, Non-official languages at work, Subsidized housing

a. Turn on the filter → Age 18 – 64;

b. Analyze → Descriptive Statistics → Crosstab → Rows= PR/CMA (enter one at a time), Column= dependent var (LICOre for example), Layer 1= IMMre. Select Chisquare from Statistics option → Note * (p<.05)/ ** (p<.01/ *** (p<.001) on the absolute difference between Imm and Non-imm.

Note: There are 5 different layers of information (themes) in CIMI 2.0: (1) Immigrant status "IMMre"; (2) Sex "SEXre"; (3) Visible minority "VISMINre"; (4) Admission category "IMMCATre"; (5) Immigrant time of arrival "YRIMre".

The instruction above is for the layer "IMMre". For admission categories and immigrant time of arrival, change IMMre to "IMMCATre"/ "YRIMre". For sex and visible minority, select Layer 1= IMMre, Layer 2= SEXre/ VISMINre.

STEP 3: FILL THE REGRESSION TABLES

1. Wages (controlled by full-time and part- time working status)

a. Turn on the filter → (AGEPfilter = 1) & (EMPL = 1) & (COWfilter = 1);

b. Run regression by going to "Analyze" → "Regression" → "Linear";

→ Dependent: sqrtWAGES

→ Independents:

— Block 1 includes 3 variables;

+ Immigrant status (IMMre)

+ Geography (for example: NF, etc.): enter one at a time

+ Immigrant status and geography variable (for example: IMMxNF)

— Block 2 includes the following variables;

+ Sex (SEXre), Age (AGECONT), Education (HCDD), Occupations (OCC_2,…10), Visible minority (VISMINre), Knowledge of official languages (Frenchre, BothEFre, NeitherEFre), Mobility status (MOB5), FULLTIME.

2. Labour force participation, Employment rate, Unemployment rate, Full-time employment rate, Non-official language at work

a. Turn on Filter → Age 18 – 64;

b. Run regression by going to "Analyze" → "Regression" → "Binary logistic";

→ Dependent: LFACTre /…(take one at a time)

→ Independents:

— Block 1 includes 3 variables;

+ Immigrant status (IMMre)

+ Geography (for example: NF, etc.): take one at a time

+ Immigrant status and geography variable (for example: IMMxNF)

— Block 2 includes the following variables;

+ Sex (SEXre), Age (AGECONT), Education (HCDD), Occupations (OCC_2,…10), Visible minority (VISMINre), Knowledge of official languages (Frenchre, BothEFre, NeitherEFre), Mobility status (MOB5)

3. LICOs, Subsidized housing

c. Turn on Filter → Age 18 – 64;

d. Run regression by going to "Analyze" → "Regression" → "Binary logistic";

→ Dependent: LICOre/ SUBSIDYre (take one at a time)

→ Independents:

— Block 1 includes 3 variables;

+ Immigrant status (IMMre)

+ Geography (for example: NF, etc.): take one at a time

+ Immigrant status and geography variable (for example: IMMxNF)

— Block 2 includes the following variables;

+ Sex (SEXre), Age (AGECONT), Education (HCDD), Occupation (NOC_2,…11), Visible minority (VISMINre), Knowledge of official languages (Frenchre, BothEFre, NeitherEFre), Mobility status (MOB5)

STEP 1: RECODE VARIABLES

(Datasets: Canadian Community Health Survey (CCHS) 2018, 2014, 2010, 2005, and 2000-01)
Note: CCHS 2018 variables will be updated when the Research Data Centre (RDC) re-opens after the COVID-19 pandemic.)

1. Filter — Age: 18 to 64

ORIGINAL VARIABLE	RECODED VARIABLE
AGE 2010 and 2014: "DHH_AGE"/ 2005: "DHHE_AGE")/ 2000-01: "DHHA_AGE" (Age groups): Continuous variable	"AGEfilter" (Age from 18-64.) 18 thru 64 → 1 = "18 – 64 years old" ELSE → 0

2. Dependent variables

ORIGINAL VARIABLE	RECODED VARIABLE
HAVE REGULAR MEDICAL DOCTOR 2010 and 2014: "HCU_1AA"/ 2005: "HCUE_1AA"/ 2000-01: TWDA_5 (Has regular medical doctor) 1 = "Yes" 2 = "No" 7 = "Don't know" 8 = "Refusal" 9 = "Not stated"	MEDIDOCTORre (Refers to the proportion of individuals who have a regular medical doctor.) 1 → 1 = Yes 2 → 0 = No 7,8,9 → Missing value
SELF-PERCEIVED LIFE STRESS 2010 and 2014: "GEN_07"/2005: GENE_07/ 2000-01: GENA_07 (Self-perceived life stress) 1 = "Not at all" 2 = "Not very" 3 = "A bit" 4 = "Quite a bit" 5 = "Extremely" 6 = "Not applicable" 7 = "Don't know" 8 = "Refusal" 9 = "Not stated"	**(1) For descriptive table:** "LIFESTRESSre" (Refers to the proportion of individuals who report being quite a bit or extremely stressed.) 4 thru 5 → 1 = "Quite a bit or extremely stressed" 1 thru 3 → 0 = "Others" 6,7,8,9 → Missing value **(2) For regression:** Use the original variable.

SELF-PERCEIVED UNMET HEALTHCARE NEEDS 2010 and 2014: "UCN_010"/ 2005: HCUE_06/ 2000-01: HCUA_06 (Unmet health care needs - self-perceived) 1 = "Yes" 2 = "No" 6 = "Not Applicable" 7 = "Don't Know" 8 = "Refusal"	UNMETHCNre (Refers to the proportion of individuals who perceive that their health care needs are not being met.) 1 → 1 = Yes 2 → 0 = No 6,7,8 → Missing value
SATISFACTION WITH LIFE IN GENERAL 2010 and 2014: "GEN_02A2"/ (Satisfaction with life in general) 0 = "Very dissatisfied" 1 = "1" 2 = "2" …. 9 = "9" 10 = "Very satisfied" 97 = "Don't know" 98 = "Refusal" 99 = "Not stated" 2005: GENE_02A Very satisfied = 1 Satisfied = 2 Neither satisfied nor dissatisfied = 3 Dissatisfied = 4 Very dissatisfied = 5 Don't know = 7 Refusal = 8 Not stated = 9 2000-01: Not available	**(1) For descriptive table:** "LIFESATISre" (Refers to the proportion of individuals who say they are very satisfied with their lives.) 9 thru 10 → 1= "Very satisfied with life" 0 thru 8 → 0 = "Others" 97,98,99 → Missing value **(2) For regression:** Use original variable. **(1) For descriptive table:** Use "LIFESATISre". "LIFESATIS" (Satisfaction with life) 5 → 1 = Very dissatisfied 4 → 2 = Dissatisfied 3 → 3 = Neither satisfied nor dissatisfied 2 → 4 = Satisfied 1 → 5 = Very satisfied 8,9 → Missing value "LIFESATISre" (Very satisfied with life) 5 → 1= "Very satisfied with life" 1thru 4 → 0 **(2) For regression:** Use LIFESATIS.

3. Key Independent Variables

ORIGINAL VARIABLE	RECODED VARIABLE
IMMIGRANT STATUS 2010 and 2014: "SDCFIMM"/ 2005: "SDCEFIMM"/ 2000-01: "SDCAFIMM" (Immigrant status) 1 = "Yes" 2 = "No" 6 = "Not applicable" 7 = "Don't know" 8 = "Refusal" 9 = "Not stated"	IMMre 1 → 1 = "Immigrants" 2 → 0 = "Non-immigrants" 6,7,8,9 → Missing value
YEAR OF IMMIGRATION 2010 and 2014: SDC_3 (Year of immigration)/ 2005: SDCE_3/ 2000-01: SDCA_3 Year 1916 - 2014 (continuous variable) Not applicable = 9996 Don't know = 9997 Refusal = 9998 Not stated = 9999	YRIMre 2009-2014 → 1 = "Recent immigrants" 1916-2008 → 2 = "Established immigrants" 997, 998, 999 → Missing value
PROVINCE 2010 and 2014: "GEO_PRV"/ 2005: GEOE_PRV/ 2000-01: GEOA_PRV (Provinces)	**(1) For descriptive table:** Do not need to recode. **(2) For regression:** **Step 1:** Create dummy variable for each province/CMA.

CMA

GEODCMAA (2014) / GEODCMA6 (2011)/ GEOEDCMA (2005) / GEOADCMA (2001)

PROVINCES	CITIES (CMAS)
Newfoundland	St. John's
Prince Edward Island	—
Nova Scotia	Halifax
New Brunswick	Moncton, Saint John
Québec	Montréal, Québec, Saguenay, Sherbrooke, Trois-Rivières
Ontario	Barrie, Brantford, Guelph, Hamilton, Kingston, Belleville, Kitchener-Cambridge-Waterloo, London, Oshawa, Ottawa-Gatineau, Peterborough, St. Catharines-Niagara, Greater Sudbury, Thunder Bay, Toronto, Windsor
Manitoba	Winnipeg
Saskatchewan	Regina, Saskatoon
Alberta	Edmonton, Calgary, Lethbridge
British Columbia	Abbotsford, Kelowna, Vancouver, Victoria

Ex: QCre (People from Quebec)

24 → 1 = "Quebec"

ELSE → 0 = "Others"

ONre (People from Ontario)

35 → 1 = "Ontario"

ELSE → 0 = "Others"

Do the same for all 10 provinces and all 35 CMAs.

Step 2: Create new variables showing the interaction between immigrant status and geographic status (Immigrant Status x Geography).

Ex: "NFxIMM", "QCxIMM" …

4. Control Variables

ORIGINAL VARIABLE	RECODED VARIABLE
VISIBLE MINORITY 2010 and 2014: "SDCDCGT"/ 2005: "SDCEGCGT"/ 2000-01: "SDCADRAC" (Cultural or racial origin - (D, G) White =1 Black =2 Korean =3 Filipino =4 Japanese =5 Chinese =6 South asian =7 Southeast asian =8 Arab =9 West asian =10 Latin american =11 Other racial or cultural origin =12 Multiple racial / cultural origins =13 Not applicable =96 Not stated =99	VISMINre (Refers to persons who are non-Caucasian in race or non-white in colour and who do not report being Aboriginal.) 1 → 0 = Not a visible minority 2-13 → 1 = Visible minority 96,99 → Missing value

SEX 2010, 2014: "DHH_SEX"/ 2005: DHHE_SEX/ 2000-01: "DHHA_SEX" (sex) 1 = "Male" 2 = "Female"	SEXre 1 → 0 = "Male" 2 → 1 = "Female"
AGE 2010, 2014: "DHH_AGE"/ 2005: "DHHE_AGE"/ 2000-01: "DHHA_SEX" (Age groups) Continuous variable	**(1) "For descriptive table:** "AGEre" (Refers to the age (in years) at last birthday before the reference date.) 18 thru 24 → 1 = "18-24 years old" 25 thru 44 → 2 = "25-44 years old" 45 thru 64 → 3 = "45-64 years old" MISSING → SYSMIS **(2) For regression:** Use the original variable.
HIGHEST LEVEL OF EDUCATION 2010, 2014: "EDUDR10" / 2005: EDUEDR10/ 2000-01: EDUADR10 Grade 8 or lower (que.Sec II or lower) = 1 Grade 9-10 (que.Sec III, IV; nfld sec I) = 2 Grade 11-13 (que. Sec V; nfld sec 2 - 3) = 3 Secondary school grad., No post-sec. = 4 Some post-secondary = 5 Trades certificate or diploma = 6 Diploma / certificate - college / cegep = 7 Univ. Certificate below bachelor's level = 8 Bachelor's degree = 9 Univ. Degree or cert. Above bac. Level = 10 Not stated = 99	**(1) For descriptive table:** EDUre (Refers to the person's most advanced certificate, diploma or degree.) 1,2,3,4,5,6,7,8 → 0= "Less than bachelor" 9,10 → 1= "Bachelor's degree or above" 99 → Missing value **(2) For regression:** Use the original variable.

KNOWLEDGE OF OFFICIAL LANGUAGES

2010, 2014: "SDC_5A_1" (Knowledge of official languages)

1 = "English only"

2 = "French only"

3 = "Both FR and EN"

4 = "Neither FR nor EN"

7 = "Don't know"

8 = "Refusal"

9 = "Not stated"

2005: Languages - can converse (SDCEDLNG)/ 2000-01: SDCADLNG

English only = 1

French only = 2

English and french only = 3

English and french and other = 4

English and other (not french) = 5

French and other (not english) = 6

Neither english nor french (other) = 7

Not stated = 99

Knowledge of Official Languages (Refers to the ability to conduct a conversation in English only, in French only, in both English and French, or in neither English or French.)

(1) For descriptive table:

2010, 2014: Don't need to recode.

2005: KOLre

1,5 → 1= English only

2,6 → 2= French only

3,4 → 3 = Both English and French

7 → 4 = Neither E nor French

99 → Missing value

(2) For regression:

Englishre (Conduct a conversation in English only)

1 → 1 = "English only"

2,3,4 → 0 = "Others"

99 → Missing value

Frenchre (Conduct a conversation in French only)

2 → 1 = "French only"

1, 3,4 → 0 = "Others"

99 → Missing value

BothEFre (Conduct a conversation in both English and French)

3 → 1 = "Both English and French"

1,2,4 → 0 = "Others"

99 → Missing value

NeitherEFre (Conduct a conversation in neither English or French)

4 → 1 = "Neither English nor French"

1thru 3 → 0 = "Others"

99 → Missing value

SELF-PERCEIVED HEALTH 2010, 2014: "GEN_01"/ 2005: GENE_01/ 2000-01: GENA_01 (Self-perceived health) 1 = "Excellent" 2 = "Very good" 3 = "Good" 4 = "Fair" 5 = "Poor" 6 = "Not applicable" 7 = "Don't know" 8 = "Refusal" 9 = "Not stated"	**(1) For descriptive table:** GENre (Refers to the percentage of people who reported that their health is very good or excellent.) 1,2 → 1 = "Very good or excellent" 3,4,5 → 0 = "Others" **(2) For regression table:** GENregre 5 → 1 = "Poor" 4 → 2 = "Fair" 3 → 3 = "Good" 2 → 4 = "Very good" 1 → 5 = "Excellent" 7,8,9 → Missing value
SELF-PERCEIVED MENTAL HEALTH 2010, 2014: "GEN_02B"/ 2005: GENE_02B/ (Self-perceived mental health) 1 = "Excellent" 2 = "Very good" 3 = "Good" 4 = "Fair" 5 = "Poor" 7 = "Don't know" 8 = "Refusal" 9 = "Not stated" 2000-01: Not available	**(1) For descriptive table:** GENMENTALre (Refers to the percentage of people who reported that their mental health is very good or excellent.) 1,2 → 1 = "very good or excellent" 3,4,5 → 0 = "others" **(2) For regression table:** GENMENTALregre 5 → 1 = "Poor" 4 → 2 = "Fair" 3 → 3 = "Good" 2 → 4 = "Very good" 1 → 5 = "Excellent" 7,8,9 → Missing value
INCOME 2010, 2014: "INC_8A"/ 2005: "INCE_4"/ 2000-01: "INCA_4" (Total pers. inc. from all sources (D,G) Continuous variable = amount in dollars 0 = None 99999996 = Not applicable 99999997 = Don't know 99999998 = Refusal 99999999 = Not stated	Income (Refers to the best estimated total personal income, before taxes and deductions, from all sources in the past 12 months.) Don't need to recode. 99999996, 99999997, 99999998, 99999999 → Missing value

OCCUPATION

2010, 2014: "LBSDOCG" (Occupation group - (D,G)

Management occupations = 01

Business, finance and admin. occ. = 02

Natural and applied sc. and rel. occ. = 03

Health occupations = 04

Occ. In social sc. / Edu. /GVT / Relig.= 05

Occ. In culture, recreation and sport = 06

Sales and service occupations = 07

Trades/trans./Equ. Operator/ rel. Occ. = 08

Occupations unique to primary industry = 09

Occ. Unique to proc. / Manuf. /Utilities = 10

Could not be coded = 95

Not applicable = 96

Not stated = 99

2005: Not available

2000-01: LBFA_31A

1 = "Management"

2 = "Professional"

3 = "Technologist"

4 = "Admin/fin/cler"

5 = "Sales/service"

6 = "Trades/transpor"

7 = "Farm/forest/fish"

8 = "Processing/manuf"

9 = "Other"

96 = "Not applicable"

97 = "Don't know"

98 = "Refusal"

99 = "Not stated"

2010, 2014:

(1) For descriptive table:
NOCre (Refers to the kind of work performed by employed persons based on the National Occupational Classification System.)

1 → 1 = Management occupations

2 thru 10 → 0 = Others

95,96,99 → Missing value

(2) For regression: NOCregre

01 thru 10 → Copy old values

96 → 11 = "Don't work"

95,99 → Missing value

Then, create dummy variable for "NOCregre": NOCregre_1, NOCregre_2,….,NOCregre_11.

2000-01:

(1) For descriptive table: NOCre

1 → 1 = Management occupations

2 thru 9 → 0 = Others

96,97,98,99 → Missing value

(2) For regression: NOCregre

1 thru 9 → Copy old values

96 → 10 = "Don't work"

97,98,99 → Missing value

Then, create dummy variable for "NOCregre": NOCregre_1, NOCregre_2,...,NOCregre_10.

FULL/PART-TIME

2010, 2014: "LBSDPFT"/ 2005: lbsedpft/2000-01: LBFADPFT (Full/Part-time - current jobs - (D))

1 = "Full-time"

2 = "Part-time"

6 = "Not applicable"

7 = "Don't know"

8 = "Refusal"

9 = "Not stated"

(1) For descriptive table: Don't need to recode

6,7,8,9 → Missing value

(2) For regression:

"FPTre" (Full/Part time)

1 → 1 = "Full-time"

2 → 2 = "Part-time"

6 → 3 = "Others" (don't work)

7,8,9 → Missing value

Then, create dummy variables:

FPTre → Full-time: 1 → 1

ELSE → 0

7,8,9 → Missing value

FPTre → Part-time: 2 → 1

ELSE → 0

7,8,9 → Missing value

FPTre → Others: 3 → 1

ELSE → 0

7,8,9 → Missing value

STEP 2: FILL THE DESCRIPTIVE TABLES & T-TEST

A. FOR CONTROL VARIABLES

1. All control variables except for "income"

a. Turn on Filter → Age 18 – 64;

b. Analyze → Descriptive Statistics → Crosstab → Rows= PR/CMA (enter one at a time), Column= all control vars, Layer 1= IMMre.

2. "Income"

a. Turn on Filter → Age 18-64;

b. Analyze → Compare mean → Dependent var: INC_8A; Layer 1: IMMre, Layer 2: PR/CMA (enter one at a time).

B. FOR INDICATORS

1. All indicators

a. Turn on Filter → Aged 18-64;

b. Analyze → Descriptive Statistics → Crosstab → Rows= PR/CMA (enter one at a time), Column= dependent var (MEDIDOCTORre, for example), Layer 1= IMMre. Select Chisquare from Statistics option → Note * (p<.05)/ ** (p<.01/ *** (p<.001) on the absolute difference between Imm and Non-imm.

Note: There are 4 different layers of information (themes) in CIMI 2.0: (1) Immigrant status "IMMre"; (2) Sex "SEXre"; (3) Visible minority "VISMINre"; (4) Immigrant time of arrival "YRIMre".

The instruction above is for the layer "IMMre". For immigrant time of arrival, change IMMre to "YRIMre". For sex and visible minority, select Layer 1= IMMre, Layer 2= SEXre/ VISMINre.

STEP 3: FILL THE REGRESSION TABLES

1. Self-perceived life stress, Satisfaction with life

a. Turn on the filter → Age 18 to 64;

b. Run regression by going to "Analyze" → "Regression" → "Linear";

→ Dependent: LIFESATISre/ LIFESTRESSre (enter one at a time)

→ Independents:

— Block 1 includes 3 variables;

+ Immigrant status (IMMre)

+ Geography (for example: NF, etc.): enter one at a time

+ Immigrant status X geography (for example: IMMxNF, etc.): take one at a time

— Block 2 includes the following variables;

+ Sex (SEXre), Age (DHH_AGE), Visible minority (VISMINre), Occupation (NOCregre), Education (EDUDR10), Income (INC_8A), Knowledge of official languages (Frenchre, BothEFre, NeitherEFre), Full-time, Self-perceived health (GENregre), Self-perceived mental health (GENMENTALregre)

2. Self-perceived unmet healthcare needs, Have a regular medical doctor

a. Turn on Filter → Age 18 – 64;

b. Run regression by going to "Analyze" → "Regression" → "binary regression";

→ Dependent: MEDIDOCTORre/ UNMETHCNre (take one at a time)

→ Independents:

— Block 1 includes 3 variables;

+ Immigrant status (IMMre)

+ Geography (for example: NF, etc.): take one at a time

+ Immigrant status X geography (for example: IMMxNF, etc.): take one at a time

— Block 2 includes the following variables (keep unchanged in all regressions);

+ Sex (SEXre), Age (DHH_AGE), Visible minority (VISMINre), Occupation (NOCregre), Education (EDUDR10), Income (INC_8A), Knowledge of official languages (Frenchre, BothEFre, NeitherEFre), Full-time, Self-perceived health (GENregre), Self-perceived mental health (GENMENTALregre)

STEP 1: RECODE VARIABLES

(Dataset: General Social Survey C27 (2013), C22 (2008), C17 (2003))

1. Filter — Age: 18 to 64

ORIGINAL VARIABLE	RECODED VARIABLE
AGE 2013, 2008, and 2003: AGE (Age of respondent at time of survey interview): Continuous variable	AGEfilter (from 18 to 64) 18 to 64 → 1 ELSE → 0

2. Dependent variables

ORIGINAL VARIABLE	RECODED VARIABLE
NUMBER OF CLOSE FRIENDS 2013: "SCF_100"/ 2008: "SCF_Q100" (Number of close friends): Continuous variable Don't know 997 Refusal 998 2003: "SCF_Q100" 1 = None 2 = 1 or 2 3= 3 to 5 4 = 6 to 10 5 = 11 to 20 6 = More than 20 8 = Not stated 9 = Don't know	Don't need to recode. Definition: Refers to an individual's average number of close friends who are not your relatives, but who you feel at ease with, can talk to about what is on your mind, or call on for help.
NUMBER OF CLOSE FRIENDS LIVING IN THE SAME CITY/COMMUNITY 2013: "SCF_102"/ 2008: "SCF_Q102" (Close friend - Same city/community) Continuous variable Valid skip = 996 Don't know = 997 Refusal = 998 2003: Not available	Don't need to recode. Definition: Refers to an individual's average number of close friends that live in the same local community or city and you feel at ease with, can talk to about what is on your mind, or call on for help.

SENSE OF BELONGING TO LOCAL COMMUNITY 2013: "SBL_100"/ 2008: "DOR_Q635"/2003: LS_Q310 (SOB to local community) 1 = Very strong 2 = Somewhat strong 3 = Somewhat weak 4 = Very weak 5 = No opinion 7 = Don't know 8 = Refusal 9 = Not stated	**(1) For descriptive table:** "BelongLCre" (Refers to the percentage of individuals who report a strong or very strong sense of belonging to their local community.) 3,4 → 0 = Very weak or somewhat weak SOB 1,2 → 1 = Very strong or somewhat strong SOB 5,7,8,9 → Missing value **(2) For regression:** "BelongLC" 4 → 1 = Very weak 3 → 2 = Somewhat weak 2 → 3 = Somewhat strong 1 → 4 = Very strong 5,7,8,9 → Missing value
SENSE OF BELONGING TO PROVINCE 2013: "SBL_300"/ 2008: "DOR_Q636"/ 2003: LS_Q320 (Sense of Belonging to Province) 1 = Very strong 2 = Somewhat strong 3 = Somewhat weak 4 = Very weak 5 = No opinion 7 = Don't know 8 = Refusal 9 = Not stated	**(1) For descriptive table:** "BelongPROVre" (Refers to the percentage of individuals who report a strong or very strong sense of belonging to their province of residence.) 3,4 → 0 = Very weak or somewhat weak SOB 1,2 → 1 = Very strong or somewhat strong SOB 5,7,8,9 → Missing value **(2) For regression:** Use "BelongPROV" 4 → 1 = Very weak 3 → 2 = Somewhat weak 2 → 3 = Somewhat strong 1 → 4 = Very strong 5,7,8,9 → Missing value

SENSE OF BELONGING TO CANADA 2013: "SBL_500"/ 2008: "DOR_Q637"/ 2003: LS_Q330 (Sense of belonging to Canada) 1 = Very strong 2 = Somewhat strong 3 = Somewhat weak 4 = Very weak 5 = No opinion 7 = Don't know 8 = Refusal 9 = Not stated	**(1) For descriptive table:** "BelongCANre" (Refers to the percentage of individuals who report a strong or very strong sense of belonging to Canada.) 3,4 → 0 = Very weak or somewhat weak SOB 1,2 → 1 = Very strong or somewhat strong SOB 5,7,8,9 → Missing value **(2) For regression:** Use "BelongCAN" 4 → 1 = Very weak 3 → 2 = Somewhat weak 2 → 3 = Somewhat strong 1 → 4 = Very strong 5 → Missing value 7,8,9 → Missing value
VICTIM OF DISCRIMINATION IN THE PAST 5 YEARS 2013: "DISCRIM" (Victim of discrimination - 5 years) Yes = 1 No = 2 Don't know = 7 Refusal = 8 Not stated = 9 2008, 2003: Not available	"DISCRIM" (Refers to the percentage of individuals who said they experienced discrimination over the past 5 years.) 1 → Yes 2 → 0 = No 7,8,9 → Missing value

3. Key independent variables

ORIGINAL VARIABLE	RECODED VARIABLE
IMMIGRANT STATUS 2013, 2008, 2003: "BRTHCAN"(Country of birth) 1 = Canada 2 = Country outside Canada 7 = Don't know 8 = Refusal 9 = Not stated	"IMMre" 1 → 0 = Non-immigrants 2 → 1 = Immigrants 7,8,9 → Missing value
YEAR OF IMMIGRATION 2013: BPR_15/ 2008: BPR_Q55/ 2003: YRARRI: Continuous variable	YRIMre (Refers to the year in which the immigrant first obtained landed immigrant or permanent resident status.) 2008-2013 → 1 = "Recent immigrants" Others → 0 = "Established immigrants" 997, 998, 999 → Missing value
ADMISSION CATEGORIES 2013: LIP_Q10 The refugee program =1 The program of re-unification with a family member alread... = 2 The points system (skilled workers and professionals, inv...= 3 Other - Specify = 4 Valid skip = 6 Don't know = 7 Refusal = 8 Not stated = 9 2008: BPR_Q60 1 = ...the refugee program? 2 = ...the program of re-unification with a family member already in Canada? 3 = ...the points system (skilled workers and professionals, investors, entrepreneurs and self-employed persons)? 4 = ...or other? 7 = Not asked 8 = Not stated 9 = Don't know 2003: Not available	IMMCATre (Refers to the name of the immigration program or group of programs under which an immigrant has been granted for the first time the right to live in Canada permanently by immigration authorities.) Don't need to recode.

PROVINCE

2013, 2008, 2003: PRV (Province of residence)

CMA

2013, 2008: STRATUM (Stratum of residence)

2003: GEO_CMA_CA (Metropolitan area and census area (code) of the respondent's residence)

PROVINCES	CITIES (CMAS)
Newfoundland	St. John's
Prince Edward Island	—
Nova Scotia	Halifax
New Brunswick	Moncton, Saint John
Québec	Montréal, Québec, Saguenay, Sherbrooke, Trois-Rivières
Ontario	Barrie, Brantford, Guelph, Hamilton, Kingston, Belleville, Kitchener-Cambridge-Waterloo, London, Oshawa, Ottawa-Gatineau, Peterborough, St. Catharines-Niagara, Greater Sudbury, Thunder Bay, Toronto, Windsor
Manitoba	Winnipeg
Saskatchewan	Regina, Saskatoon
Alberta	Edmonton, Calgary, Lethbridge
British Columbia	Abbotsford, Kelowna, Vancouver, Victoria

(1) For descriptive table: Do not need to recode.

(2) For regression:

Step 1: Create dummy variable for each province/CMA.

Ex: QCre (People from Quebec)

24 → 1 = "Quebec"

ELSE → 0 = "Others"

ONre (People from Ontario)

35 → 1 = "Ontario"

ElSE → 0 = "Others"

Do the same for all 10 provinces and all 35 CMAs.

Step 2: Create new variables showing the interaction between immigrant status and geographic status (Immigrant Status x Geography).

Ex: "NFxIMM", "QCxIMM" …

4. Control variables

ORIGINAL VARIABLE	RECODED VARIABLE
VISIBLE MINORITY 2013, 2008: "VISMIN" (Visible minority status of the respondent) 1 = "Visible minority" 2 = "Not a visible minority" 6 = "Valid skip" 7 = "Don't know" 8 = "Refusal" 9 = "Not stated" 2003: Not available	VISMINre (The Employment Equity Act defines visible minorities as "persons, other than Aboriginal peoples, who are non-Caucasian in race or non-white in colour".) 1 → 1 = "Visible minority" 2 → 0 = "Not a visible minority" 6,7,8,9 0 → Missing value

AGE 2013, 2008, 2003: AGE (age of respondent at time of survey interview): Continuous variable	**(1) For descriptive table:** "AGEre" (CIMI Age groups) 18 thru 24 → 1 = "18-24 years old" 25 thru 44 → 2 = "25-44 years old" 45 thru 64 → 3 = "45-64 years old" MISSING → SYSMIS **(2) For regression:** Use the original variable.
SEX 2013, 2008, 2003: SEX (Sex of respondent) 1 = "Male" 2 = "Female"	SEXre 1 → 0 = "Male" 2 → 1 ="Female"
HIGHEST LEVEL OF EDUCATION 2013: EHG1_01 (Education - Highest degree) 1 = Less than high school diploma or its equivalent 2 = High school diploma or a high school equivalency certificate 3 = Trade certificate or diploma 4 = College/CEGEP/other non-university certificate or diploma 5 = University certificate or diploma below the bachelor's level 6 = Bachelor's degree (e.g. B.A., B.Sc., LL.B.) 7 = University certificate, diploma, degree above the BA level Valid skip = 96 Don't know = 97 Refusal = 98 Not stated = 99 2008, 2003: EDU10 (Highest level education obtained by respondent (10 groups) 1 = Doctorate/masters/some graduate 2 = Bachelor's degree 3 = Diploma/certificate from community college 4 = Diploma/certificate from trade/technical 5 = Some university 6 = Some community college/cegep/nursing 7 = Some trade/technical 8 = High school diploma 9 = Some secondary/high school 10 = Elementary school/no schooling 98 = Not stated 99 = Don't know	Definition: Refers to the person's most advanced certificate, diploma or degree. 2013: **(1) For descriptive table:** EDUre 1,2,3,4,5 → 0= "Less than bachelor" 6,7 → 1= "Bachelor's degree or above" 96, 97, 98, 99 → Missing value **(2) For regression:** Use the original variable 96, 97, 98, 99 → Missing value 2008, 2003: **(1) For descriptive table:** EDUre 3,4,5 ,..10 → 0 = "Less than bachelor" 1,2 → 1 = "Bachelor's degree or above" 98, 99 → Missing value **(2) For regression:** EDU 10 → 1 = Elementary school/No schooling 9 → 2 = Some secondary/High school ... 2 → 9 = Bachelor's degree 1 → 10 = Doctorate/Masters/Some graduate 98, 99 → Missing value

MOTHER TONGUE LANGUAGE

2013:

LNR_111 (Still understand first childhood language - English)

LNR_112 (Still understand first childhood language -French)

LNR_113...26 (Still understand first childhood language- other languages)

1 = Yes, still understood

2 = No

7 = Not asked

8 = Not stated

2008, 2003:

"LANCHSUE" (Still understand first childhood language - English)

"LANCHSUF" (Still understand first childhood language - French)

"LANCHSUO" (Still understand first childhood language - others)

1 = Yes, still understood

2 = No

7 = Not asked

8 = Not stated

Definition: Refers to the first language learned at home in childhood and still understood by the person at the time the data was collected.

The variable we need for the analysis should include 4 categories: English only, French only, Both English & French, Neither English nor French.

LNR_111/LANCHSUE → ENGMT

1 → 1

2 → 0

7,8 → Missing value

LNR_112/LANCHSUF → FREMT

1 → 1

2 → 0

7,8 → Missing value

LNR_113/ LANCHSUO → NeitherEFre

1 → 1

2 → 0

7,8 → Missing value

BothEFre (Both English and French)

Compute → Transform: BothEFre = (ENGMT=1) & (FREMT=1)

Note: Don't define missing values for BothEFre. If we do that, all of the missing values of ENGMT & FREMT will be excluded, which leads to the problem of losing many valid samples in Englishre and Frenchre.

Englishre

Compute → Transform: Englishre = (ENGMT=1) &(BothEF=0)

Frenchre

Compute → Transform: Frenchre = (FREMT=1) & (BothEF=0)

→ Four variables we need are Englishre, Frenchre, BothEFre, NeitherEFre, which are used for both descriptive and regression tables.

INCOME 2013 and 2008: "INR_032"/ 2003: IN_Q0120 (Personal income): continuous variable Valid skip = 99999996 Don't know = 99999997 Refusal = 99999998 Not stated = 99999999	Definition: Refers to the best estimated total personal income, before taxes and deductions, from all sources in the past 12 months. Don't need to recode. 99999996,7,8,9 → Missing value
FULL-TIME WORKERS 2013, 2008, 2003: WKWEHR (Number of hours usually worked at all jobs in a week) Continuous variable 997 = Not asked 998 = Not stated 999 = Don't know	**(1) For descriptive table:** FullTimeRe (Full-time employment consists of persons who usually work 30 hours or more per week.) 1,2,...,29 → 0 = "Part-time" 30,.. → 1 = "Full-time" 997,998,999 → Missing value **(2) For regression:** Full-time 1,2,...,29 → 0 = "Others" 30,.. → 1 = "Full-time" 997 → 0 = "Others" 998,999 → Missing value
OCCUPATION 2013: Not available 2008: NOCS2006_C10 (National Occupational Classification (2006) of the respondent - last 12 months - 10 categories) 2003: SOC91C10 (Standard Occupational Classification (1991) of the respondent - 10 categories) 1 = Management occupations 2 = Business, finance and administrative occupations 3 = Natural and applied sciences and related occupations 4 = Health occupations 5 = Occupations in social science, education, government service and religion 6 = Occupations in art, culture, recreation and sport 7 = Sales and services occupations 8 = Trades, transport and equipment operators and related occupations 9 = Occupations unique to primary industry 10 = Occupations unique to processing, manufacturing and utilities 96 = Not applicable 97 = Not asked 98 = Not stated 99 = Don't know	**(1) For descriptive table:** NOCre (Refers to the kind of work performed by employed persons based on the National Occupational Classification System) 1 → 1 = Management occupations 2,3,4,..,10 → 0 = Others 96, 97,98,99 → Missing value **(2) For regression:** NOCregre 1,2,3,...,10 → copy old values 96 → 11 = "Not applicable/ Don't work" 97,98,99 → Missing value

STEP 2: FILL THE DESCRIPTIVE TABLES & T-TEST

A. FOR CONTROL VARIABLES

1. All control variables except for "income"

a. Turn on Filter → Age 18 – 64;

b. Analyze → Descriptive Statistics → Crosstab → Rows= PR/CMA (enter one at a time), Column= all control all control variables, Layer 1= IMMre.

2. "Income"

a. Turn on Filter → Age 18-64;

b. Analyze → Compare mean → Dependent var: INC_032; Layer 1: IMMre, Layer 2: PR/CMA (enter one at a time).

B. FOR INDICATORS

1. Number of close friends, Number of close friends living in the same city/community

a. Turn on Filter → Aged 18-64;

b. Analyze → Compare mean → Dependent var: SCF_100/ SCF_200; Layer 1: IMMre, Layer 2: PR/CMA (enter one at a time);

c. For t-test (don't change the filter) — 2 steps;

* Split file – Group based on PR/CMA – Oke

* Analyze → Compare means → independent sample t-test: test var – SCF_100/ SCF_200; Group variable: IMMre (define g1: 1 (Imm); g2: 2 (Non-imm) → take p value from independent sample test → t-test for equality of means (Sign(2-tailed)

→ Note * (p<.05)/ ** (p<.01/ *** (p<.001) on the absolute difference between Imm and Non-imm.

d. Turn off "split file" function.

2. Sense of belonging to local community/Province/Canada, Victim of discrimination

a. Turn on Filter – Aged 18 – 64;

b. Analyze → Descriptive Statistics → Crosstab → Rows= PR/CMA (enter one at a time), Column= DISCRIM/BelongLCre/...; Layer 1= IMMre. Select Chisquare from Statistics option. → Note * (p<.05)/ ** (p<.01/ *** (p<.001) on the absolute difference between Imm and Non-imm.

Note: There are 5 different layers of information (themes) in CIMI 2.0: (1) Immigrant status "IMMre"; (2) Sex "SEXre"; (3) Visible minority "VISMINre"; (4) Admission category "IMMCATre"; (5) Immigrant time of arrival "YRIMre".

The instruction above is for the layer "IMMre". For admission categories and immigrant time of arrival, change IMMre to "IMMCATre"/ "YRIMre". For sex and visible minority, select Layer 1= IMMre, Layer 2= SEXre/ VISMINre.

STEP 3: FILL THE REGRESSION TABLES

1. Number of close friends, number of close friends living in the same city/community

a. Turn on the filter → Age 18 to 64;

b. Run regression by going to "Analyze" → "Regression" → "Linear";

→ Dependent: SCF_100/ SCF_200

→ Independents:

— Block 1 includes 3 variables;

+ Immigrant status (IMMre)

+ Geography (for example: ONre, etc.): take one at a time

+ Immigrant status X geography (for example: IMMxON, etc.): take one at a time

— Block 2 includes the following variables (keep unchanged in all regressions);

+ Sex (SEXre), Age(AGE), Visible minority(VISMINre), Occupation (NOCregre), Education (EDU/ EHG1_01), Income (INC_032), Mother tongue language (FRENCHre, BothEFre, NeitherEFre), FullTime

2. Victim of discrimination, sense of belonging to local community/province/Canada

a. Turn on Filter → Age 18 – 64;

b. Run regression by going to "Analyze" → "Regression" → "Binary regression";

→ Dependent: DISCRIM/BelongLC/…

→ Independents:

— Block 1 includes 3 variables;

+ Immigrant status (IMMre)

+ Geography (for example: NF, etc.): take one at a time

+ Immigrant status X geography (for example: IMMxNF, etc.): take one at a time

— Block 2 includes the following variables;

+ Sex (SEXre), Age(AGE), Visible minority(VISMINre), Occupation (NOCregre), Education (EDU/ EHG1_01), Income (INC_032), Mother tongue language (FRENCHre, BothEFre, NeitherEFre), FullTime

STEP 1: RECODE VARIABLES

(Dataset: General Social Survey C27 (2013), C22 (2008), C17 (2003)

1. Filter — Age: 18 to 64

ORIGINAL VARIABLE	RECODED VARIABLE
AGE 2013, 2008, and 2003: AGE (Age of respondent at time of survey interview): Continuous variable	AGEfilter (from 18 to 64) 18 to 64 → 1 ELSE → 0

2. Dependent variables

ORIGINAL VARIABLE	RECODED VARIABLE
UNPAID VOLUNTEER WORK IN THE PAST 12 MONTHS 2013, 2008: VCG_300/ 2003: VCG_Q300 (Volunteer work - past 12 months) Yes = 1 No = 2 Don't know = 7 Refusal = 8	"VOLUNTre" (Refers to the percentage of individuals who have done unpaid volunteer work in the past 12 months.) 1 → 1 = "Yes" 2 → 0 = "No" 7,8 → Missing value
INVOLVEMENT IN ORGANIZATIONS IN THE PAST 12 MONTHS 2013: CERD230/2008: CER_Q150/ 2003: CE_Q260 (Number of organization types - 12 months) Continuous variable	"InvolveORG" (Refers to the percentage of individuals who reported membership, participation or involvement in groups or organizations in the past 12 months.) 0 → 0 = "No" ELSE → 1 = "YES" 97,98,99 → Missing value
VOTED IN THE LAST PROVINCIAL ELECTION 2013: VBR_30/ 2008: PER_Q120/ 2003: PE_Q120 (Last provincial election – voted) Yes = 1 No = 2 Valid skip = 6 Don't know = 7 Refusal = 8 Not stated = 9	"VotedPROV" (Refers to the proportion of individuals who voted in the last provincial election.) 1 → 1 = Yes (Voted) 2 → 0 = No (Did not vote) 3, 6,8,9 → Missing value

VOTED IN THE LAST FEDERAL ELECTION 2013: VBR_Q10/ 2008: PER_Q110/ 2003: PE_Q110 (Last federal election – Voted) Yes = 1 No = 2 Valid skip = 6 Don't know = 7 Refusal = 8 Not stated = 9	"VotedFED" (Refers to the proportion of individuals who voted in the last federal election.) 1 → 1 = Yes (Voted) 2 → 0 = No (Did not vote) 6,7,8,9 → Missing value

3. Key independent variables

ORIGINAL VARIABLE	RECODED VARIABLE
IMMIGRANT STATUS 2013, 2008, 2003: "BRTHCAN"(Country of birth) 1 = Canada 2 = Country outside Canada 7 = Don't know 8 = Refusal 9 = Not stated	"IMMre" 1 → 0 = Non-immigrants 2 → 1 = Immigrants 7,8,9 → Missing value
YEAR OF IMMIGRATION 2013: BPR_15/ 2008: BPR_Q55/ 2003: YRARRI: Continuous variable	YRIMre (Refers to the year in which the immigrant first obtained landed immigrant or permanent resident status.) 2008-2013 → 1 = "Recent immigrants" Others → 0 = "Established immigrants" 997, 998, 999 → Missing value

ADMISSION CATEGORIES 2013: LIP_Q10 The refugee program =1 The program of re-unification with a family member alread... = 2 The points system (skilled workers and professionals, inv...= 3 Other - Specify = 4 Valid skip = 6 Don't know = 7 Refusal = 8 Not stated = 9 2008: BPR_Q60 1 = ...the refugee program? 2 = ...the program of re-unification with a family member already in Canada? 3 = ...the points system (skilled workers and professionals, investors, entrepreneurs and self-employed persons)? 4 = ...or other? 7 = Not asked 8 = Not stated 9 = Don't know 2003: Not available	IMMCATre (Refers to the name of the immigration program or group of programs under which an immigrant has been granted for the first time the right to live in Canada permanently by immigration authorities.) Don't need to recode.

PROVINCE

2013, 2008, 2003: PRV (Province of residence)

CMA

2013, 2008: STRATUM (Stratum of residence)

2003: GEO_CMA_CA (Metropolitan area and census area (code) of the respondent's residence)

PROVINCES	CITIES (CMAS)
Newfoundland	St. John's
Prince Edward Island	—
Nova Scotia	Halifax
New Brunswick	Moncton, Saint John
Québec	Montréal, Québec, Saguenay, Sherbrooke, Trois-Rivières
Ontario	Barrie, Brantford, Guelph, Hamilton, Kingston, Belleville, Kitchener-Cambridge-Waterloo, London, Oshawa, Ottawa-Gatineau, Peterborough, St. Catharines-Niagara, Greater Sudbury, Thunder Bay, Toronto, Windsor
Manitoba	Winnipeg
Saskatchewan	Regina, Saskatoon
Alberta	Edmonton, Calgary, Lethbridge
British Columbia	Abbotsford, Kelowna, Vancouver, Victoria

(1) For descriptive table: Do not need to recode.

(2) For regression:

Step 1: Create dummy variable for each province/CMA.

Ex: QCre (People from Quebec)

24 → 1 = "Quebec"

ELSE → 0 = "Others"

ONre (People from Ontario)

35 → 1 = "Ontario"

ElSE → 0 = "Others"

Do the same for all 10 provinces and all 35 CMAs.

Step 2: Create new variables showing the interaction between immigrant status and geography (Immigrant Status X Geography).

Ex: "NFxIMM", "QCxIMM" …

4. Control variables

ORIGINAL VARIABLE	RECODED VARIABLE
VISIBLE MINORITY 2013, 2008: "VISMIN" (Visible minority status of the respondent) 1 = "Visible minority" 2 = "Not a visible minority" 6 = "Valid skip" 7 = "Don't know" 8 = "Refusal" 9 = "Not stated" 2003: Not available	VISMINre (The *Employment Equity Act* defines visible minorities as "persons, other than Aboriginal peoples, who are non-Caucasian in race or non-white in colour".) 1 → 1 = "Visible minority" 2 → 0 = "Not a visible minority" 6,7,8,9 → Missing value
AGE 2013, 2008, 2003: AGE (Age of respondent at time of survey interview): Continuous variable	**(1) For descriptive table:** "AGEre" (CIMI Age groups) 18 thru 24 → 1 = "18-24 years old" 25 thru 44 → 2 = "25-44 years old" 45 thru 64 → 3 = "45-64 years old" MISSING → SYSMIS **(2) For regression:** Use the original variable.
SEX 2013, 2008, 2003: SEX (Sex of respondent) 1 = "Male" 2 = "Female"	SEXre 1 → 0 = "Male" 2 → 1 = "Female"

HIGHEST LEVEL OF EDUCATION

2013: EHG1_01 (Education - highest degree)

1 = Less than high school diploma or its equivalent

2 = High school diploma or a high school equivalency certificate

3 = Trade certificate or diploma

4 = College/CEGEP/other non-university certificate or diploma

5 = University certificate or diploma below the bachelor's level

6 = Bachelor's degree (e.g. B.A., B.Sc., LL.B.)

7 = University certificate, diploma, degree above the BA level

Valid skip = 96

Don't know = 97

Refusal = 98

Not stated = 99

2008, 2003: EDU10 (Highest level education obtained by respondent (10 groups)

1 = Doctorate/masters/some graduate

2 = Bachelor's degree

3 = Diploma/certificate from community college

4 = Diploma/certificate from trade/technical

5 = Some university

6 = Some community college/cegep/nursing

7 = Some trade/technical

8 = High school diploma

9 = Some secondary/high school

10 = Elementary school/no schooling

98 = Not stated

99 = Don't know

Definition: Refers to the person's most advanced certificate, diploma or degree.

2013:

(1) For descriptive table: EDUre

1,2,3,4,5 → 0= "Less than bachelor"

6,7 → 1= "Bachelor's degree or above"

96, 97, 98, 99 → Missing value

(2) For regression: Use the original variable.

96, 97, 98, 99 → Missing value

2008, 2003:

(3) For descriptive table: EDUre

3,4,5 ,..10 → 0 = "Less than bachelor"

1,2 → 1 = "Bachelor's degree or above"

98, 99 → Missing value

(4) For regression: EDU

10 → 1 = Elementary school/No schooling

9 → 2 = Some secondary/High school

...

2 → 9 = Bachelor's degree

1 → 10 = Doctorate/Masters/Some graduate

98, 99 → Missing value

MOTHER TONGUE LANGUAGE

2013:

LNR_111 (Still understand first childhood language - English)

LNR_112 (Still understand first childhood language - French)

LNR_113…26 (Still understand first childhood language - other languages)

1 = Yes, still understood

2 = No

7 = Not asked

8 = Not stated

2008, 2003:

"LANCHSUE" (Still understand first childhood language - English)

"LANCHSUF" (Still understand first childhood language - French)

"LANCHSUO" (Still understand first childhood language - other languages)

1 = Yes, still understood

2 = No

7 = Not asked

8 = Not stated

Definition: Refers to the first language learned at home in childhood and still understood by the person at the time the data was collected.

The variable we need for the analysis should include 4 categories: English only, French only, Both English & French, Neither English nor French.

LNR_111/LANCHSUE → ENGMT

1 → 1

2 → 0

7,8 → Missing value

LNR_112/LANCHSUF → FREMT

1 → 1

2 → 0

7,8 → Missing value

LNR_113/ LANCHSUO → NeitherEFre

1 → 1

2 → 0

7,8 → Missing value

BothEFre (Both English and French)

Compute → Transform: BothEFre = (ENGMT=1) & (FREMT=1)

Note: Don't define missing values for BothEFre. If we do that, all of the missing values of ENGMT & FREMT will be excluded, which leads to the problem of losing many valid samples in Englishre and Frenchre.

Englishre

Compute → Transform: Englishre = (ENGMT=1) & (BothEF=0)

Frenchre

Compute → Transform: Frenchre = (FREMT=1) & (BothEF=0)

→ Four variables we need are Englishre, Frenchre, BothEFre, NeitherEFre, which are used for both descriptive and regression tables.

INCOME 2013 and 2008: "INR_032"/ 2003: IN_Q0120 (Personal income): continuous variable Valid skip = 99999996 Don't know = 99999997 Refusal = 99999998 Not stated = 99999999	Definition: Refers to the best estimated total personal income, before taxes and deductions, from all sources in the past 12 months. Don't need to recode. 99999996,7,8,9 → Missing value
FULL-TIME 2013, 2008, 2003: WKWEHR (Number of hours usually worked at all jobs in a week) Continuous variable 997 = Not asked 998 = Not stated 999 = Don't know	**(1) For descriptive table:** FullTimeRe (Full-time employment consists of persons who usually work 30 hours or more per week) 1,2,...,29 → 0 = "Part-time" 30,.. → 1 = "Full-time" 997,998,999 → Missing value **(2) For regression:** FullTime 1,2,...,29 → 0 = "Others" 30,.. → 1 = "Full-time" 997 → 0 = "Others" 998,999 → Missing value
OCCUPATION 2013: Not available 2008: NOCS2006_C10 (National Occupational Classification (2006) of the respondent - last 12 months - 10 categories) 2003: SOC91C10 (Standard Occupational Classification (1991) of the respondent - 10 categories) 1 = Management occupations 2 = Business, finance and administrative occupations 3 = Natural and applied sciences and related occupations 4 = Health occupations 5 = Occupations in social science, education, government service and religion 6 = Occupations in art, culture, recreation and sport 7 = Sales and services occupations 8 = Trades, transport and equipment operators and related occupations 9 = Occupations unique to primary industry 10 = Occupations unique to processing, manufacturing and utilities 96 = Not applicable 97 = Not asked 98 = Not stated 99 = Don't know	**(1) For descriptive table:** NOCre (Refers to the kind of work performed by employed persons based on the National Occupational Classification System.) 1 → 1 = Management occupations 2,3,4,...,10 → 0 = Others 96, 97,98,99 → Missing value **(2) For regression:** NOCregre 1,2,3,...,10 → copy old values 96 → 11 = "Not applicable/ Don't work" 97,98,99 → Missing value

STEP 2: FILL THE DESCRIPTIVE TABLES & T-TEST

A. FOR CONTROL VARIABLES

1. All control variables except for "income"

a. Turn on Filter → Age 18 – 64;

b. Analyze → Descriptive Statistics → Crosstab → Rows= PR/CMA (enter one at a time), Column= all control variables, Layer 1= IMMre.

2. "Income"

a. Turn on Filter → Age 18-64;

b. Analyze → Compare mean → Dependent var: INC_032; Layer 1: IMMre, Layer 2: PR/CMA (enter one at a time).

B. FOR INDICATORS

1. All indicators

a. Turn on Filter → Aged 18-64;

b. Analyze → Descriptive Statistics → Crosstab → Rows= PR/CMA (enter one at a time), Column= dependent var (MEDIDOCTORre, for example), Layer 1= IMMre. Select Chisquare from Statistics option → Note * (p<.05)/ ** (p<.01/ *** (p<.001) on the absolute difference between Imm and Non-imm.

Note: There are 5 different layers of information (themes) in CIMI 2.0: (1) Immigrant status "IMMre"; (2) Sex "SEXre"; (3) Visible minority "VISMINre"; (4) Admission category "IMMCATre"; (5) Immigrant time of arrival "YRIMre".

The instruction above is for the layer "IMMre". For admission categories and immigrant time of arrival, change IMMre to "IMMCATre"/ "YRIMre". For sex and visible minority, select Layer 1= IMMre, Layer 2= SEXre/ VISMINre.

STEP 3: FILL THE REGRESSION TABLES

1. For all indicators

a. Turn on Filter → Aged 18-64;

b. Run regression by going to "Analyze" → "Regression" → "Binary regression";

→ Dependent: VOLUNTre/…

→ Independents:

— Block 1 includes 3 variables;

+ Immigrant status (IMMre)

+ Geography (for example: NF, etc.): take one at a time

+ Immigrant status X geography (for example: IMMxNF): take one at a time

— Block 2 includes the following variables;

+ Sex (SEXre), Age(AGE), Visible minority(VISMINre), Occupation (NOCregre), Education (EDU/ EHG1_01), Income (INC_032), Mother tongue language (FRENCHre, BothEFre, NeitherEFre), FullTime