

Building a Data Story from Census



Rong Luo, PhD
December 14, 2022

Outline

- Introduction
- Understanding Data Set
- Building a Data Story from Census
 - An Example

➤ Introduction

The Census is sent to all Canadian households every 5 years.

- **Short Form:** Every household receives
 - Date of birth and age
 - Sex
 - Relationships with household members (ex: married or common-law)

- Knowledge of official languages
- Language spoken most often
- Other language spoken regularly
- First language spoken



➤ Introduction

▪ **Long Form:** 25% of Canadian households

• Activities of daily living, such as location of work, mode of transportation to work, and other related information

- Sociocultural information, such as ethnicity and visible minority information
- Mobility
- Place of birth
- Education
- Labor market activities
- Housing

Combined with income data from the tax returns of Canadians and immigration data from the Immigration, Refugees and Citizenship Canada immigration database



➤ Understanding Data Set

- Microdata
- Individual Response Data
- Census
- The Canadian Population
- Public Use Microdata File (PUMF)
- Sensitive Information Removed





➤ Understanding Data Set

■ Data Access

leddy.uwindsor.ca/ad

Here you go!

<odesi>

leddy.uwindsor.ca/ad

Search

- Library Catalogue
- Journal Articles & Research Tools
- Journal List
- Order From Other Libraries
- Course Readings

Help

- Research Help
- Statistics and Data Help
- Writing Help
- Computer Help
- Services
- Borrowing
- Select Course Resource Lists
- Contact Us
- About Ledy

Home

Academic Data Centre

Please be advised that starting Tuesday, March 27, drop-in data, GIS and statistical support will be available remotely via email to align with the campus emergency measures to minimize COVID-19 risk. Email us and we will be responding between 12:30PM-4:30PM Monday-Friday. Thank you for your patience during this time.

Librarians and data experts at the ADC are available to researchers and students for consultation. Get expert help on finding datasets, accessing confidential Statistics Canada data, reviewing statistical methods and procedures, creating graphs and maps, or managing your research data. Email us for an appointment.

Get started with data by reading one of our research guides.

Key sources of data

Statistics Canada
Data from the nation's statistical agency—the starting point for any statistical research on Canada. Access Census and other survey data. Alternatively, access StatCan data through CANSIM @ OASIS, for easier bulk-download.

ICPSR
One of the world's largest archives of social science data, ICPSR includes social surveys, polls, longitudinal data files, and more, from countries around the world.

re3data.org
A registry of research data repositories. Use this to search for over 2,000 research data repositories, which store primary research data deposited by researchers and research groups across all disciplines.

UWindsor Dataserve
The University of Windsor's institutional data repository. Search for primary research data created by Windsor researchers.

Data and statistics research guides

- Business and Financial Data
- Canadian Census Data
- Computer Science Data
- Data on Developing Countries
- Education Statistics and Data
- Environmental, Engineering, Pollution and Climate Data

Academic Data Centre Services

- Statistical software and services
- Research data management
- UWindsor Dataserve
- Statistics Canada Research Data Centre

Geospatial services

- Getting started
- Geospatial data guides
- GIS help and tutorials
- GIS software
- GIS news and blogs

Tweets by @UWindsorADC

The Academic Data Centre (@UWindsorADC) is in response to the COVID-19 pandemic, the 8th Canadian Statistical Student Conference will be online (May 30) & open to the public. Registration for the event is now FREE! docs.google.com/forms/d/1FAAL...

Registration - It... in response to the... [docs.google.com](https://docs.google.com...)

Apr 14, 2020

The Academic Data Centre (@UWindsorADC) For those conducting COVID-19 health research: Emergency use ICD codes for COVID-19 disease outbreak (as of March 23) www.ecdc.europa.eu/en/...

Embed View on Twitter

➤ Understanding Data Set

- List of Some Variables





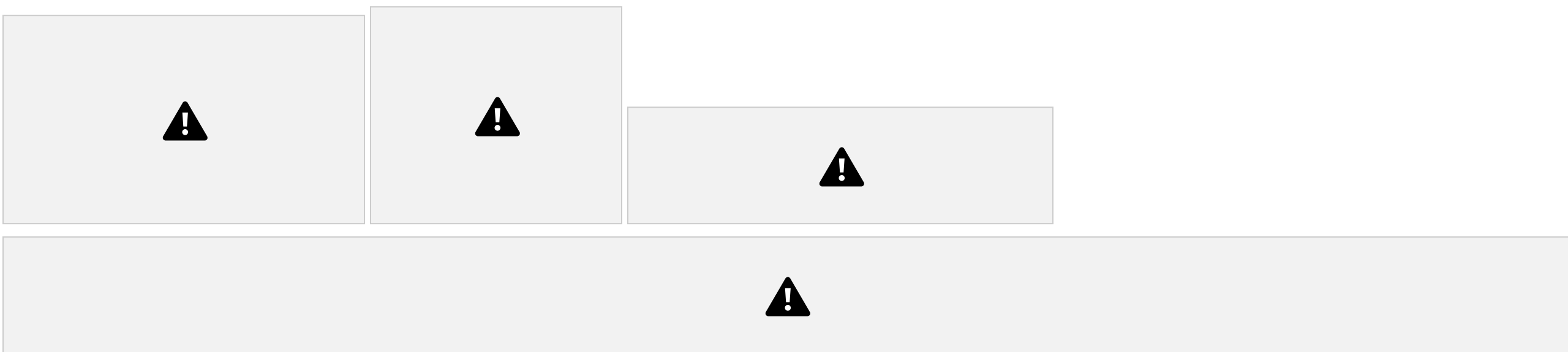
➤ Understanding Data Set

- Total 123 Variables – 2016 Census

Demographic, social, and economic information of Canadian

- Excellent for Researchers Performing Statistical Analysis

Relationships between variables can be analyzed



➤ Building a Data Story from Census

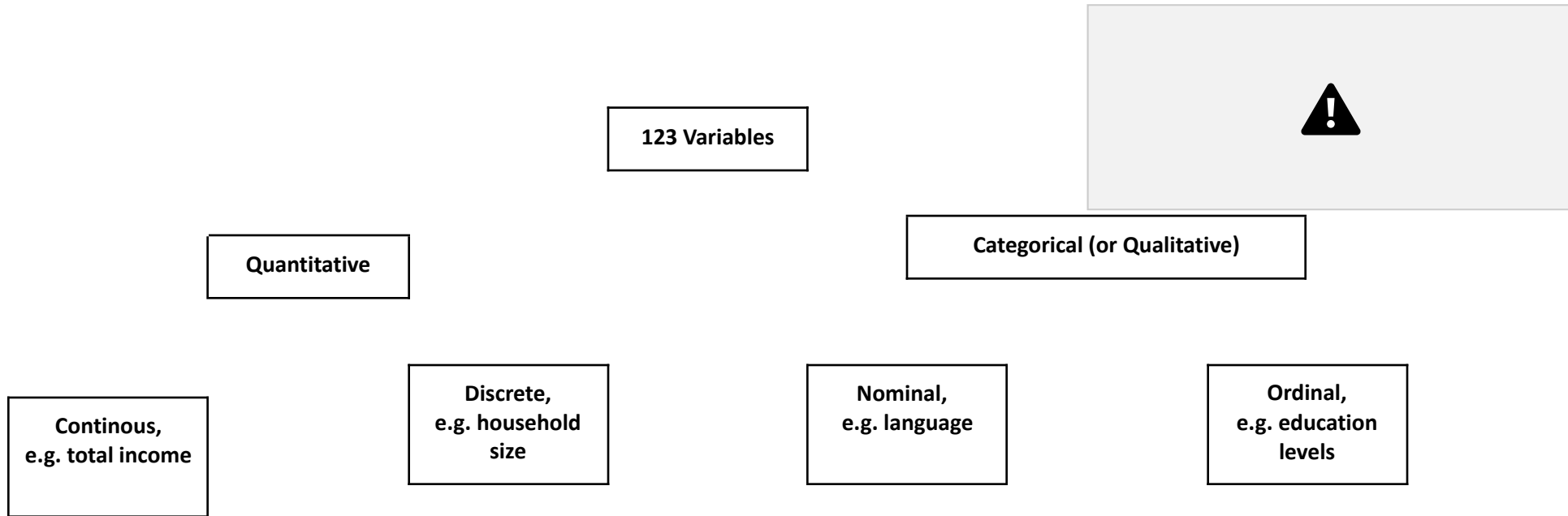


Figure 1. Type of Variables (Finlay & Agresti, 1986)



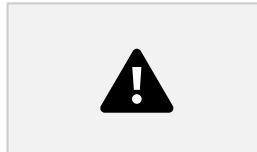
➤ Building a Data Story from Census

- Descriptive Statistics

Table 1. Descriptive Statistics

Types of Variables Statistics Graphs

Quantitative mean, and SD



Qualitative frequency and percentage



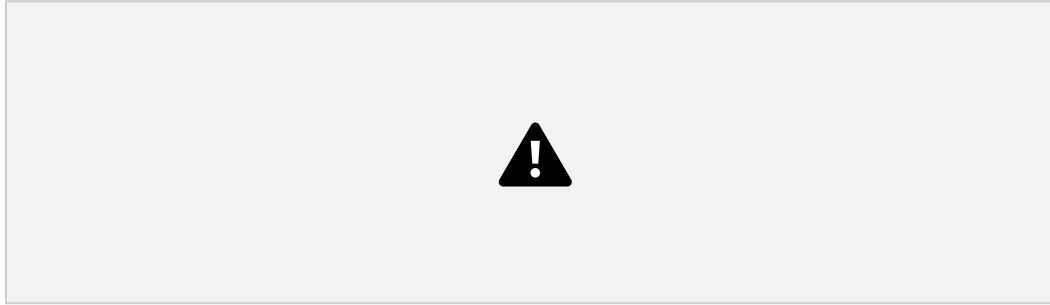


➤ Building a Data Story from Census

- Statistical Tests

- **T-test:** Two groups mean (e.g., comparing mean income of male and female)
- **ANOVA:** More than two groups mean (e.g., comparing mean income among people in Windsor, London, and Toronto. Post hoc test will be required)
- **Correlation:** Linear relationship between two continuous variables (e.g., total income with shelter cost)

- **Chi-square:** The association between two qualitative variables (e.g., detecting association between low income status and knowledge of official language)



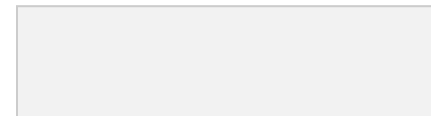
➤ **Building a Data Story from Census**

- **Example**

- *Research Question?*

Does education background affect the total income of Ontario senior immigrant?

- *Data?*





➤ **Building a Data Story from Census**

- Example - Continued
- Descriptive Statistics: Total Income (Dollars)

Table 2. Descriptive Statistics of Total Income of Ontario Senior Immigrants



Figure 2. Distribution of Total Income of Ontario Senior Immigrants

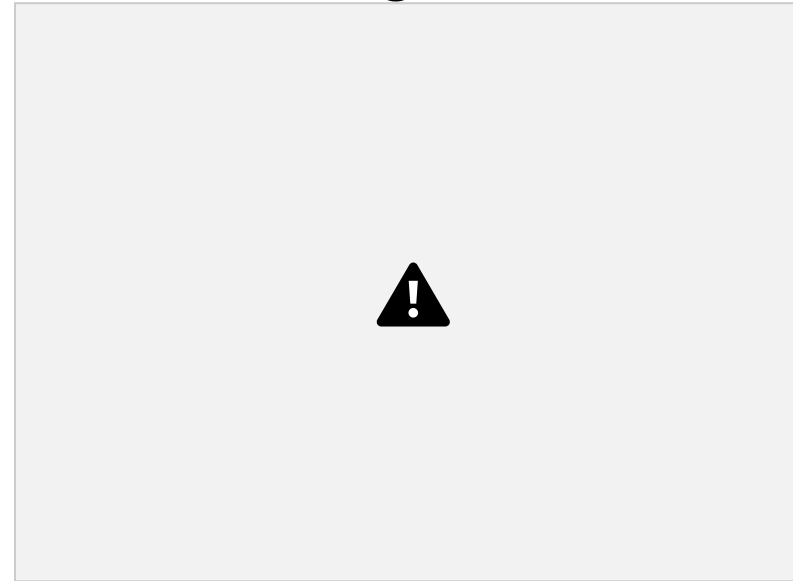


➤ **Building a Data Story from Census**

- Example - Continued

- Descriptive Statistics: Education of Ontario Senior Immigrants

Table 3. Education of Ontario Senior Immigrants




		

Figure 3. A Pie Chart of the Data in Table 3



➤ Building a Data Story from Census

- Example - Continued

ANOVA showed that the effect of education was significant, $p < 0.001$.

Post hoc analyses indicated that the mean income of people with high school diploma or below have lower income than people with some college or university education .





THANK YOU
ANY QUESTIONS?

