

# Survey Data: Design and Examples

M. Ehsan Karim

2022-09-10



# Contents

|          |                                     |           |
|----------|-------------------------------------|-----------|
| <b>1</b> | <b>Outline</b>                      | <b>5</b>  |
| <b>2</b> | <b>Model-based Approach</b>         | <b>7</b>  |
| 2.1      | Example . . . . .                   | 7         |
| 2.2      | Research question . . . . .         | 7         |
| 2.3      | Data . . . . .                      | 8         |
| 2.4      | Checking assumptions . . . . .      | 11        |
| 2.5      | Analysis . . . . .                  | 13        |
| 2.6      | Verdict . . . . .                   | 19        |
| 2.7      | Exercises (Optional) . . . . .      | 20        |
| <b>3</b> | <b>Design-based Approach</b>        | <b>21</b> |
| 3.1      | Sampling . . . . .                  | 21        |
| 3.2      | Statistical inference . . . . .     | 23        |
| 3.3      | Complex surveys . . . . .           | 25        |
| 3.4      | Other ideas . . . . .               | 28        |
| 3.5      | Further readings . . . . .          | 28        |
| 3.6      | Exercise . . . . .                  | 28        |
| <b>4</b> | <b>Potential Data Sources</b>       | <b>29</b> |
| 4.1      | Survey data with features . . . . . | 29        |
| 4.2      | Others . . . . .                    | 29        |

|          |   |           |
|----------|---|-----------|
| <b>5</b> | <b>Importing CCHS to R</b>                | <b>31</b> |
| 5.1      | Downloading CCHS data from UBC . . . . .  | 31        |
| 5.2      | Reading and Formatting the data . . . . . | 38        |
| 5.3      | Processing data in R . . . . .            | 54        |
| <b>6</b> | <b>Demystifying NHANES</b>                | <b>57</b> |
| 6.1      | Overview . . . . .                        | 57        |
| 6.2      | Survey history . . . . .                  | 57        |
| 6.3      | NHANES datafile and documents . . . . .   | 58        |
| 6.4      | Exercise (web) . . . . .                  | 61        |
| <b>7</b> | <b>Importing NHANES to R</b>              | <b>63</b> |
| 7.1      | NHANES Dataset . . . . .                  | 63        |
| 7.2      | Accessing NHANES Data . . . . .           | 63        |

# Preamble

## Description

Surveys are designed to gather and assess national level health-related information. Usually governments collect these data, and make some effort to publish a version of the data for researchers to use. In this document, we will talk about various survey designs, some examples, how to download some of these data sources from Canadian and US sources. An example is included at the end regarding how to reproduce results from a paper that used one of these open data sources.

## Outline

- Review of Model-based Approach
- Introduction to Design-based Approach
- Complex survey design examples
- Canadian Community Health Survey - Annual Component (CCHS)
  - Data import to R
- National Health and Nutrition Examination Survey (NHANES)
  - Understanding NHANES data and documentation structure
  - Data import to R
  - Reproducing results from an article

## Version history

Materials were updated over time through various deliveries of the content:

- Teaching materials for SPPH 504-007 (till 2021)
- Teaching materials for SPPH 604 (2022-)