



## Introduction to Data Analysis

Week 5 - Tutorial

### Contents

1	Instructions	2
2	Exercise 1: Introduction to Missing values [Est. time: 10 min]	2
3	Exercise 2: Summarise and Tabulate Missing values [Est. time: 20 min]	2
4	Exercise 3: Visualise Missing values [Est. time: 20 min]	2
5	Exercise 4: Exploring Melbourne Housing data [Est. time: 25 min]	3

## Workshop objectives

- ✓ Introduction to Missing values
- ✓ Summarise and Tabulate Missing values
- ✓ Visualise Missing values
- ✓ Representing missing values in a Tidy way

### 1 Instructions

1. Follow this link [here](#) and log in into Rstudio Cloud as explained in the workshop video in Moodle.
2. In each question you will replace '\_\_\_\_' with your answer. Knit the document once you've answered all the questions. Remember to go to `knitr::opts_chunk` at line 9, change `eval = TRUE`, then knit the document.

### 2 Exercise 1: Introduction to Missing values [Est. time: 10 min]

1. Once you have logged in into the cloud space, open `Ex1_intro.Rmd`
2. This exercise has theory and some practice examples.
3. Read the Theory as a pre-workshop step.

### 3 Exercise 2: Summarise and Tabulate Missing values [Est. time: 20 min]

1. Open `Ex2_summarise-missings.Rmd`.
2. We will explore dataset `riskfactors` from package `naniar`.
3. This file is split into 2 sections, A and B.
  - Section A: Summarise Missing values [Est. time: 10 min]
  - Section B: Tabulate Missing values [Est. time: 10 min]
4. Read instructions and answer the questions for both A and B.

### 4 Exercise 3: Visualise Missing values [Est. time: 20 min]

1. Open `Ex3_visualise-missings.Rmd`.
2. This file is split into 2 sections, A and B.
  - Section A: Visualizing All of missingness [Est. time: 10 min]
  - Section B: Visualising missing cases and variables [Est. time: 10 min]
3. Read instructions and answer the questions for both A and B.
4. There are some extra bits with other datasets. Don't spend any time on it during the Workshop hours. You can try them and explore more in your own time, if interested.

## 5 Exercise 4: Exploring Melbourne Housing data [Est. time: 25 min]

1. Open `Ex4_housing.Rmd`.
2. Answer the questions Q1-Q7.