Course 2 Section 1.13 - Missing values

Jiaying Wu

02/10/2020

```r
# library
library(tidyverse)
library(visdat)
library(naniar)

# Read the tb data and filter for Australia only
tb_au <- read_rds(here::here("data", "tb_tidy.rds")) %>%
  filter(country == "Australia")
```

Q1: What proportion of values are missing?

```r
miss_summary(tb_au)["miss_df_prop"]
```

```
## # A tibble: 1 x 1
##   miss_df_prop
##          <dbl>
## 1       0.0417
```

Q2: What proportion of variables have missing values?

```r
miss_summary(tb_au)["miss_var_prop"]
```

```
## # A tibble: 1 x 1
##   miss_var_prop
##           <dbl>
## 1         0.333
```

Q3: How many cases have two missing values?

```r
miss_case_table(tb_au)
```

```
## # A tibble: 3 x 3
##   n_miss_in_case n_cases pct_cases
##            <int>   <int>     <dbl>
## 1              0     192      85.7
## 2              1       8      3.57
## 3              2      24      10.7
```